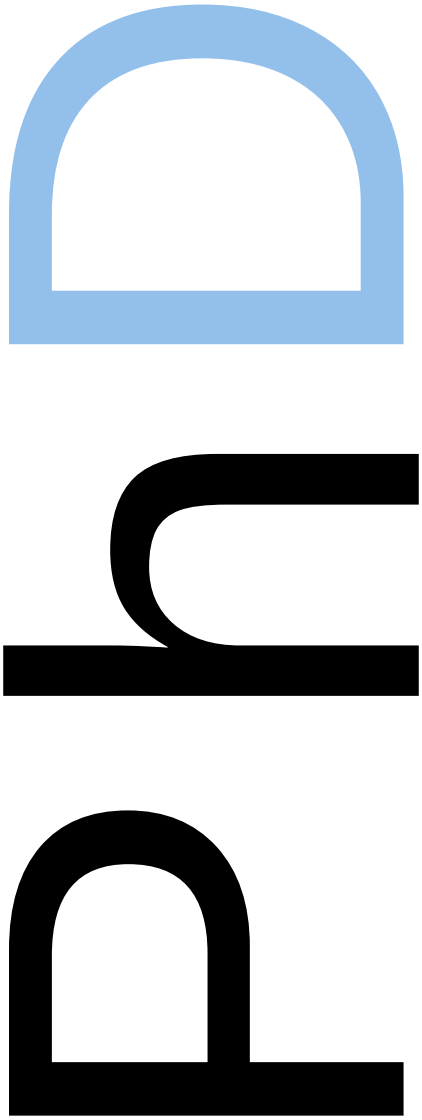
Nonlinear optical conductivity of crystals in the independent electron approximation

Daniel Passos
Doutoramento em Física
Departamento de Física e Astronomia
2021

Orientador
João Manuel Borregana Lopes dos Santos, Professor Catedrático,
Faculdade de Ciências, Departamento de Física e Astronomia



Nonlinear optical conductivity of crystals in the independent electron approximation

Faculty of Sciences, University of Porto

D. J. Passos

Publications by the author

The research in this thesis was done in collaboration with Gonçalo Ventura, João Lopes dos Santos and João Viana Parente Lopes. From our fruitful discussions, several publications resulted that influenced this dissertation. Chapters 4 and 5 of this thesis, in particular, are heavily based on the following publications:

- "Nonlinear optical responses of crystalline systems: Results from a velocity gauge analysis", D. J. Passos, G. B. Ventura, J. M. B. Lopes dos Santos, J. Viana Parente Lopes, N. M. R. Peres, *Physical Review B* (2018)
- "Nonlinear optical conductivity of a two-band crystal I.", D. J. Passos, G. B. Ventura, J. M. B. Lopes dos Santos, J. Viana Parente Lopes, *Journal of Physics: Condensed Matter* (2021)

respectively.

Other publications that I co-authored and relate to the subject of, but do not feature prominently in, this thesis are:

- "Gauge covariances and nonlinear optical responses", G. B. Ventura, D. J. Passos, J. M. B. Lopes dos Santos, J. Viana Parente Lopes, N. M. R. Peres, *Physical Review B* (2017)
- "A study of the nonlinear optical response of the plain graphene and gapped graphene monolayers beyond the Dirac approximation", G. B. Ventura, D. J. Passos, J. Viana Parente Lopes, J. M. B. Lopes dos Santos, *Journal of Physics: Condensed Matter* (2020)
- "Second order divergence in the third order DC response of a cold semiconductor", G. B. Ventura, D. J. Passos, J. Viana Parente Lopes, J. M. B. Lopes dos Santos, *arXiv:2004.01919* (2020)
- Comment on "Jerk current: a novel bulk photovoltaic effect", G. B. Ventura, D. J. Passos, J. Viana Parente Lopes, J. M. B. Lopes dos Santos, *Physical Review Letters* (2021)

Unrelated to the topic of the thesis and involving previous collaborations, but published during the course of the PhD, are the following works:

- "On the ferroelectric and magnetoelectric mechanisms in low Fe^{3+} doped $TbMnO_3$", Rui Vilarinho, Eugénia Queirós, D. J. Passos et al., *Journal of Magnetism and Magnetic Materials* (2017)
- "Suppression of the cooperative Jahn-Teller distortion and its effect on the Raman octahedra-rotation modes of $TbMn_{1-x}Fe_xO_3$", Rui Vilarinho, D. J. Passos, E. Queirós et al., *Physical Review B* (2018)

Acknowledgements

First of all, I would like to give my most sincere thanks to my supervisor, João Lopes dos Santos, for the opportunity to pursue a PhD on theoretical condensed matter physics. If I stayed in Porto for my doctoral studies, it was due to his acceptance to guide my research efforts. Even before, he had served as a source of inspiration and knowledge, for its excellent lecturing, introducing me to the subject of condensed matter physics itself, and for the example he represented, as a researcher and as a professor. During the course of developing my research, I benefited immensely from our discussions and debates and from his remarkable physical insight. But most of all, I have to thank him for the freedom I was given to pursue my own ideas, even when they steered us away from the planned path, and for his patience.

As far as supervision is concerned, I must extend my thanks to João Viana Parente Lopes, who often acted as a *de facto* co-supervisor, and with which I shared many conversations, on a variety of topics not limited to research, and from which I learned a lot. I must also acknowledge him for his invaluable help when either code wasn't running properly or some other technical issue had me stuck, as well as for teaching me simple, but extremely useful, things like ssh connections and other "computator stuff" of which I was (and largely still am) fairly ignorant. Importantly, I must recognize him for his readiness and availability to intervene whenever an urgent problem or inopportune crisis emerged, for which I am most thankful.

Completing my group of collaborators of the past years is Gonçalo Ventura, without which this thesis would not exist, or would certainly look very, or should I say, completely different. The work that fills these pages was the result of our countless discussions, exchanges of ideas and a joint travel through a series of promising avenues, that seemed to lead to insurmountable obstacles the next day, and whose difficulties were already contradicted and resolved the day after. For his collaboration and support in navigating the troubled waters of theoretical physics research, and for his friendship, I am thankful.

I'm thankful also for all the colleagues I had the opportunity to interact with, from Porto and Minho, during the course of the Map-Fis programme, and for the friends I made and with which I shared not only office hours, but lunch conversations and much non-physics related fun.

Finally, I would like to thank my family, for the support they provided all my life and specially during this last pandemic year.

Abstract

In this thesis, the theory of the electronic nonlinear optical conductivity of crystals is explored under the independent electron and electron dipole approximations, with relaxation considered at a phenomenological level.

The problem is treated in a semiclassical fashion: a quantum system of electrons moving in the periodic potential of the crystal is coupled to a classical optical field. The nonlinear conductivity is derived by means of a perturbative expansion of the density matrix in powers of the optical field. Two equivalent formulations of the perturbation theory are commonly used: the so-called velocity gauge and length gauge methods. These methods are analyzed in detail here and new insights are provided.

In the velocity gauge, a minimal coupling Hamiltonian is adopted and the optical field couples only states with the same Bloch vector in the first Brillouin zone. The treatment presented in this work differs from the one in the literature, as it is developed in more general grounds: it retains its validity for finite band models and circumvents the difficulties that have been traditionally associated with this method. This formulation is well suited for numerical calculations, where it permits nonlinear conductivities to be computed for any frequency, beyond the usual low energy descriptions, without added difficulty.

The standard length gauge method, in its current form, dates back to the early nineties. We revisit this approach and show that a clearer and more insightful use of it can be made by having the nonlinear conductivity broken up into fundamental, easily calculable, pieces, based on the possible resonances between optical frequencies and band energies. Analogously to how a Fermi golden rule calculation gives the linear optical response, the nonlinear conductivity can be obtained by evaluation of a small number of integrals.

Regardless of the method used to derive it, the final expression for the nonlinear conductivity is constrained by symmetry. Beyond the usual tensor relations derived from specific crystal symmetries, general statements can be made on the basis of overall permutation symmetry and time-reversal symmetry. Overall permutation symmetry is defined here for complex frequencies over the entire Argand plane, making it applicable even for lossy media. The signatures and consequences of these two important symmetries are identified.

As a demonstration of the principles described here, the linear, second and third order conductivity of monolayer graphene are studied, with the possibility of a substrate induced gap.

Contents

Chapter 1

Introduction

This thesis is entirely devoted to the study of the nonlinear optical properties of solids, as encoded in a specific set of response functions: the nonlinear optical conductivities. These are the central objects in perturbative treatments of nonlinear optics and are responsible for a broad range of phenomena, from harmonic generation to the intensity dependent refractive index.

Schemes for computing these quantities were devised in the early sixties based on standard perturbation theory. At the time, the focus was on the nonlinear optical properties of dielectric inorganic solids and dilute media such as atomic gases. For crystals, the complications brought by the existence of energy bands and the extended character of the Bloch functions delayed progress until the nineties, when clear and applicable formulas for the nonlinear conductivity were finally obtained.

Nonetheless, a survey of the literature will show such calculations to be involved and opaque. It is my goal here to clarify the structure of these response functions in the case of crystals and present methods by which they can be straightforwardly computed. After the historic overview that will soon follow and a presentation of the standard perturbation theory in Chapter 2, two new methods are described in succeeding chapters.

The first method, presented in Chapter 4, is entirely new and resolves a decades long problem concerning equivalent gauge choices. It is, in a sense, a reformulation of the minimal coupling approach that extends its generality and frees it from any unphysical infrared divergences.

The second method is given in Chapter 5 and is more of a close look at the widely adopted perturbation theory of Aversa and Sipe that attempts to bring a greater clarity to this theory by inspection of the existing resonances in the nonlinear conductivity. It is shown that besides insight, considerable simplification of the calculations can be achieved by the reasoning introduced here.

These methods provide complementary and improved ways to compute the nonlinear conductivity of crystals, both numerically and analytically. They are demonstrated by calculations of the nonlinear optical response of a series of two-band models. These models are first presented in Chapter 3.

Further understanding of the structure of the nonlinear conductivity follows from symmetry constraints. The usual symmetries are reviewed and focus is placed on (a generalized version of) overall permutation symmetry. The implications of these symmetries, or their absence, are addressed in Chapter 6 of this work.

Chapter 7 summarizes the findings and discusses topics for future research.

The subject of nonlinearities in condensed media is a complex one and some simplifying assumptions must be made. The discussion presented here will be confined to the electronic contribution to the nonlinear conductivity of crystals, typically dominant at optical frequencies. Relaxation of the generated currents is introduced via a phenomenological parameter (Section 2.5).

Additionally, two fundamental approximations are assumed throughout this thesis: the independent electron and the electric dipole approximations. The latter consists of neglecting the spatial dispersion of the nonlinear conductivity and is reasonable for most cases of interest, since the wavelength of light is usually much greater than the unit cell size. Conversely, the neglect of electron-electron interactions can be considered a drastic oversimplification. Still, attending to the complexity of the nonlinear response of solids, it seems recommendable to first attain a proper understanding of the response at the single electron level. Also, comparisons with experiment suggest that, at least for gapless systems, qualitative agreement can be found.

There is a certain elegance and simplicity in the structure of the nonlinear conductivity that is easily lost beneath the formal weight of second and third order perturbation theory and it is the purpose of this text to bring them forth.

1.1 Constitutive relation

The central equations of electrodynamics were discovered in the nineteen century, first written down by Maxwell [1], and later rephrased by Heaviside [2] using his vector calculus.

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \tag{1.1}$$
$$\nabla \cdot \mathbf{B} = 0 \tag{1.2}$$
$$\nabla \times \mathbf{E} = -\partial_t \mathbf{B} \tag{1.3}$$
$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \, \partial_t \mathbf{E} + \mu_0 \, \partial_t \mathbf{J} \tag{1.4}$$

Here, the fields $\mathbf{E}$ and $\mathbf{B}$ will be regarded as classical.

Considering the significant impact the field theory of electromagnetism had in physics, and society in general, one could be inclined to think that the study of light, an electromagnetic phenomenon, ought to have been exhausted by now. Yet, optics, the branch of physics that explores the behavior and uses of light is ever evolving, with new discoveries and inventions put forth every year. Part of this incessant exploration is no doubt due to the richness of the underlying field theory and the technological implications that stem from a greater control over the properties of light. But equally significant is the diversity of scenarios that nature presents us by means of a wide variety of materials through which light can propagate or be absorbed.

When a light wave moves through any material, it affects the motion of the charged particles therein, be they free carriers or bound charges. This charge motion builds up to a macroscopic current density $\mathbf{J}$ which acts as a source of new electromagnetic waves. If the relation between the current density and the optical

fields, termed the constitutive relation[1], is specified, then Eqs. 1.1-1.4 can be solved for the optical field $\mathbf{E}(t)$, allowing for a complete description of the evolution of any light pulse in the medium.

The constitutive relation $\mathbf{J}(\mathbf{E})$ can be rather complicated even for simple systems. In attempting to formulate a description of some generality, it is natural to first consider the limit of weak optical fields, sufficiently weak to merit a power series expansion,

$$\mathbf{J}(t) = \mathbf{J}^{(1)}(t) + \mathbf{J}^{(2)}(t) + \cdots + \mathbf{J}^{(n)}(t) + \ldots \tag{1.5}$$

The first term in this expansion is a linear combination of the optical fields, evaluated at any given time.

$$J^{\beta\,(1)}(t) = \int_{-\infty}^{+\infty} \sigma^{\beta\alpha}(t,t')\, E^{\alpha}(t')\, dt' \tag{1.6}$$

with an implicit summation over repeated tensor indices α. The coefficients are our response functions. This expression is however too general; some symmetries are always present that constrain this linear relation. From causality, it follows that $\sigma^{\beta\alpha}(t,t') = 0$ for $t' > t$, and time translation symmetry[2] implies [3]

$$J^{\beta\,(1)}(t) = \int_{-\infty}^{+\infty} \sigma^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' \tag{1.7}$$

This is the usual definition of the optical conductivity. Similarly, we can take higher powers of $\boldsymbol{E}$ into account,

$$J^{\beta\,(n)}(t) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sigma^{\beta\alpha_1\ldots\alpha_n}(t-t_1,\ldots,t-t_n)\, E^{\alpha_1}(t_1)\ldots E^{\alpha_n}(t_n)\, dt_1\ldots dt_n \tag{1.8}$$

This equation defines the nonlinear conductivity of order n. Specifying the form of the nonlinear conductivities for a medium equates to deriving its constitutive relation.

Eq. 1.8 is written in the time-domain. Often, the light pulse is well defined in frequency (e.g. CW lasers) and it is then more useful to express the constitutive relation in the frequency domain,

$$J^{\beta\,(n)}(t) = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi}\ldots\frac{d\omega_n}{2\pi}\; \sigma^{\beta\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n)\, E^{\alpha_1}(\omega_1)\ldots E^{\alpha_n}(\omega_n)\, e^{-i(\omega_1+\ldots+\omega_n)t} \tag{1.9}$$

with

$$E^{\alpha_i}(\omega_i) \equiv \int_{-\infty}^{+\infty} E^{\alpha_i}(t)\, e^{i\omega_i t}\, dt \tag{1.10}$$

[1] Actually, it is more common to have the constitutive relation defined as the functional relationship between the electric polarization and the optical fields. The response function is then the susceptibility instead of the conductivity - the connection between the two is explored in Appendix B. However, there are several reasons for why it is preferable to consider the electric current instead of the polarization, especially in the presence of free carriers.

[2] Here it is meant that the entire system is to be translated, the field and the medium; the optical experiment could be carried out at any time of the day, without it affecting the results.

and

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n) \equiv \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}\sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n)\,e^{i\omega_1 t_1}\ldots e^{i\omega_n t_n}\,dt_1\ldots dt_n \tag{1.11}$$

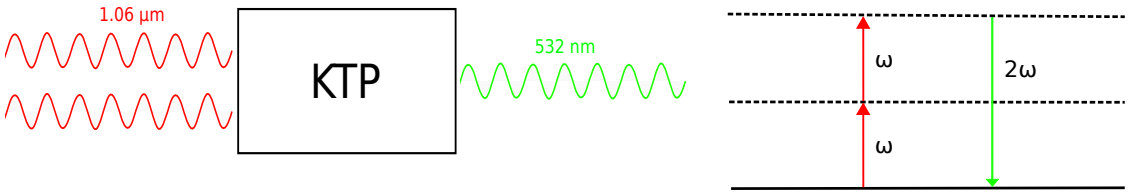
For the Fourier transforms in Eq. 1.11 to converge, the time-domain conductivity must vanish for $t_i \to +\infty$. In other words, the current is required to decay when the optical field is removed in order for the frequency domain description to be sensible. We will return to this point later (see Section 2.5). The convergence for negative times is ensured by causality.

In most circumstances, the first term in Eq. 1.5 will suffice. This is the regime of linear optics, that we observe in our everyday experience. In it, a monochromatic light wave incident on a transparent material will propagate at a medium-dependent speed, whilst maintaining its frequency of oscillation. Stating it in the languages of particles, if a light beam comprised of red photons is sent into the material, the photons that shoot out on the other side will still be red photons. A monochromatic beam may refract, diffract, experience birefringence and other linear effects, but retains its frequency (color). If a secondary monochromatic beam is used, the first beam is unaffected by its presence. Optical waves do not interact with each other, nor do they significantly alter the optical properties of the medium they propagate in.

These commonplaces no longer hold when the constitutive relation becomes nonlinear. The properties of the medium are then dependent on the intensity of illumination and optical waves experience medium-mediated interactions.

The reason a linear relation can often be assumed at optical frequencies is that the optical fields are usually much weaker than the microscopic atomic fields that bind the electrons. If coherent intense light is used however, as that provided by lasers, then the higher order terms can no longer be neglected.

In these situations, the lowest order non-zero nonlinear term in Eq. 1.5 is often sufficient to describe the light-matter interaction. This means that most consideration is given to the second or third order conductivity. If the medium lacks inversion symmetry, then the main focus lies on the second order conductivity, while in centrosymmetric media all even orders in Eq. 1.5 vanish and the third order conductivity is of greater interest.



(a) (b)

Figure 1.1: Second harmonic generation: (a) A light beam at a wavelength $2\pi c/\omega = 1.06\,\mu$m (infrared) incident on a KTP crystal is converted into radiation at $2\pi c/(2\omega) = 532\,$nm (green light); (b) Energy-level diagram of the medium illustrating the absorption of two infrared photons and emission of a green photon. This process is commonly used in green laser pointers [4], where green laser light is generated from infrared lasers.

Various nonlinear phenomena become possible due to the second and third order conductivities. If we consider two light waves at frequencies ω_1 and ω_2 (for convenience, let us suppose $\omega_1 > 2\omega_2$), it can be gleaned from Eq. 1.9 that at second order current oscillations will be induced at $\omega_1 + \omega_2$ (sum frequency generation - SFG), $\omega_1 - \omega_2$ (difference frequency generation - DFG), $2\omega_1$ and $2\omega_2$ (second harmonic generation - SHG). In the third order, more possibilities arise: $2\omega_1 + \omega_2$ (SFG); $\omega_1 + 2\omega_2$ (SFG); $2\omega_1 - \omega_2$ (DFG); $\omega_1 - 2\omega_2$ (DFG); $3\omega_1$ and $3\omega_2$ (third harmonic generation - THG), among others. Crucially, the current is a source term in Maxwell's equations and these oscillating currents will generate optical waves at these new optical frequencies. It is then possible to have mixing of light waves to generate new waves of distinct frequencies. The case of second harmonic generation is illustrated in Fig. 1.1.

Many special cases exist, such as those involving DC fields (take for instance ω_2=0 in the examples above), DC currents (the so-called shift [5], injection [5] and jerk [6–8] currents) and the effects a light wave has on itself due to the changes it causes in the medium it propagates (optical Kerr effect, intensity-dependent refractive index, self-phase modulation). These effects will not be discussed in this thesis in any detail, as the emphasis lies in providing a framework that works for any set of optical frequencies. Some of the main second and third order nonlinear effects can be seen in Table 1.1, together with the nonlinear conductivities that describe them. More appropriate and extensive discussions can be found in several standard textbooks on the subject [3, 9, 10]. A useful resource letter written by Garmire [11] compiles introductory references to the many specific topics encompassed by nonlinear optics.

Nonlinear conductivity	Nonlinear optical effects
$\sigma(\omega,\omega)$	Second harmonic generation
$\sigma(\omega,0)$	Linear electro-optic (Pockels) effect
$\sigma(\omega,-\omega)$	Photogalvanic effect; shift and injection currents
$\sigma(\omega,0,0)$	Quadratic electro-optic (DC Kerr) effect
$\sigma(\omega,\omega,\omega)$	Third harmonic generation
$\sigma(\omega,\omega,0)$	Field-induced second harmonic generation
$\sigma(\omega,-\omega,\omega)$	Optical Kerr effect; intensity dependent refractive index; self-phase modulation; self-focusing
$\sigma(\omega,-\Omega,\Omega)$	Optical Kerr effect; intensity dependent refractive index; cross-phase modulation
$\sigma(\omega,\omega,-2\omega)$	Two-color current injection
$\sigma(\omega,-\omega,0)$	Jerk, injection and shift currents

Table 1.1: Different frequency components of the nonlinear conductivity describe distinct nonlinear optical effects. Specific references on most of these effects can be found in [11].

1.2 A historical overview

In 1961, Franken, Hill, Peters and Weinreich reported the first observation of optical second harmonic generation [12]. At the Harrison Randall Laboratory of Physics in Michigan, they had a beam of coherent red light, emitted by a ruby laser, traverse a piece of crystalline quartz and form a saturated dark spot in a spectographic plate standing on the other side. A smaller spot formed about thirty-five centimeters to its left. That tiny spot was caused by the arrival of a few ultraviolet photons generated in the quartz crystal and marked the birth of a new subfield of optics: nonlinear optics.

Their discovery would not have been possible without Maiman's invention of the laser the previous year [13, 14]. It was the emergence of this new source of intense coherent radiation (with field strengths up to the order of 10^7 V/m) that was to propel a series of fundamental discoveries over the incoming decade on new nonlinear optical phenomena [15], of which Franken's *et al* observation of harmonic generation was the first[3]. Conversely, nonlinear optics diversified available laser technology and expanded its potential applications [4].

In the same month the work on second harmonic generation was presented, another group reported the first observation of two-photon absorption [16]. This was quickly followed by some theoretical order of magnitude estimations on the two-photon absorption cross-section [17] (the possibility of multiphoton absorption was first conceived and analyzed by Goeppert-Mayer in her PhD thesis, way back in 1931). Two-photon absorption is directly related to the real part of the third order conductivity [3].

Optical mixing of a more general character, including sum-frequency generation [18] (observed by setting two ruby lasers at different temperatures), optical rectification [19] and DC-field induced second harmonic generation [20], among others, was explored during the early sixties. It was realized very early that much higher conversion efficiencies in harmonic generation were possible if one ensured the involved waves were *phase-matched*. Phase-matching of optical waves was achieved by proper use of birefringent crystals [21, 22] (e.g. KDP). All these nonlinear effects were encompassed in the theoretical framework set by Armstrong *et al* in 1962 [23]. This work included a derivation of the nonlinear response functions based on perturbation theory as well as an extensive discussion on the set of coupled wave equations that resulted when the nonlinear constitutive relation was inserted in Maxwell's equations. The solutions of these wave equations describe optical mixing and propagation effects in nonlinear media [3, 10].

The importance of symmetry was not left unnoticed in these early investigations. At the time, a pertinent question was if the measured optical nonlinearities were due to ionic or electronic motions (or both?). This was answered with a suggestion by Kleinman [24] to take a closer look to the symmetry properties of the susceptibility (or conductivity) tensors. Considerations of the crystal's point symmetry had in fact been used to discard the possibility of artifact in Franken's *et al* experiment [12]. It was known that the symmetry of the second order susceptibility (conductivity) was

[3]From the effects in Table 1.1, some of those involving DC fields form an exception as they do not necessarily require intense laser light. The DC Kerr and Pockels effects were observed in the second half of the nineteen century, much before the birth of the laser and are often presented as precursors to the development of nonlinear optics [15].

identical to that of piezoelectric tensors, which had already undergone proper classification for the various crystal symmetry groups [24, 25]. Kleinman pointed out that the measurements that had been made of harmonic generation used optical frequencies well below those of electronic transitions and well above ionic ones, therefore the dispersion of the nonlinear tensors should be negligible if the nonlinearities were due to electronic contributions and not so otherwise. The dispersionless tensors have a greater symmetry and hence a smaller number of independent components, which could be tested experimentally. It was soon confirmed that the electronic contributions were in fact dominant [26].

Kleinman symmetry is a special case of overall permutation symmetry [3, 10]. This more general symmetry was demonstrated in [23] for lossless systems. An extension of the validity of overall permutation symmetry was proved in [27] by extending the domain region of the nonlinear conductivity to the complex plane (see Section 2.5). The most interesting aspect of this symmetry is that it connects different frequency components of the nonlinear conductivity, thereby providing a link between distinct nonlinear effects. An example that was verified early on is the connection between optical rectification and the linear electro-optic effect in the "clamped lattice" regime [19]. Another consequence of this symmetry are the Manley-Rowe relations [27–29] that relate the emitted and absorbed powers at different optical frequencies in lossless nonlinear media. Overall permutation symmetry is discussed extensively in Chapter 6.

Nonlinear conductivities successfully provide us with a unified understanding of nonlinear optics in the perturbative regime, but the actual calculation of the nonlinear conductivity for a specific medium is challenging [3]. The cumbersome expressions that follow directly from perturbation theory often come accompanied with some subtle difficulties and require (sometimes considerable) analytical work to be put in a form that can be of use. Most importantly, a requirement to accurately evaluate such expressions is a detailed knowledge of the perturbation matrix elements and the electronic energies of the system [30].

The ability to compute the second and third order conductivities is of great value in determining the materials of interest for applications and in the design of optical devices involving nonlinear processes [3, 31]. It can also be useful in retrieving information about the materials themselves [4, 31]. The nonlinear optical properties of crystals are known to be more sensitive to the details of the wavefunctions and eigenergies than their linear counterparts, making the accuracy of such calculations a more stringent test on the validity of band structure theories [30, 32].

During the seventies, the focus was on atomic and molecular gases. For sufficiently simple molecules, *ab initio* calculations of the dipole matrix elements were possible and found good agreement with experiment [3, 33, 34]. This was the case for harmonic generation in alkali-metal vapours [34], with their simple hydrogen-like spectra, and for molecular hydrogen [33], where calculations of the third order susceptibility matched the measured values of Raman gain and of the spontaneous Raman cross-section. For more complex systems, semiempirical methods were adopted with measured values of transition energies and oscillator strengths incorporated into the calculation [3]. A pioneering effort in this front were the works on p-nitroaniline [4]

[4]Organic compounds like p-nitroaniline with highly charge-correlated π electron states attracted great interest at the time due to their unusually high optical nonlinearities. This makes it perhaps less surprising that graphene was later shown to have a strong nonlinear optical response.

that used well established molecular orbital calculations to obtain quantitative predictions for the second order susceptibility[5] [35–38]. These ultimately reproduced the experimental data rather well [38].

Comprehensive studies of the nonlinear optical response of crystals came much later. The first measurements showing the dispersion of the second order conductivity were presented in [39]. Some correlation was found between critical points of the band structure and pronounced features in the nonlinear optical response. Theoretically, the first proper full band structure calculation was performed by Fong and Shen [40] on zincblende semiconductors using an empirical pseudopotential approach to predict the behavior of the nonlinear conductivity over an extended frequency range[6]. It is worth emphasizing that one of their conclusions was that the $\boldsymbol{k}$-dependence of the dipole matrix elements could not be ignored[7]. A decade later, their work was expanded on by a research group from University of Toronto that presented their own full band structure calculations on the nonlinear optical response of semiconductors [5, 30, 32, 41–44], using empirical tight binding models and, later, density functional theory. Their works included comparisons between different methods of obtaining the perturbation matrix elements [30], a study on the anisotropy in the third order response of silicon [41] and, in 1990, the first calculations of the third order conductivities of semiconductors over an extended frequency range [32]. That it took decades from the development of the general framework for solids [25] to its application to specific systems is a testament to the complexity of the task.

Most of the works discussed so far used a minimal coupling Hamiltonian,

$$\hat{H} = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\mathbf{A}(t)) \tag{1.12}$$

where H_0 is the unperturbed Hamiltonian of the crystal and $\mathbf{A}(t)$ is the vector potential: $\mathbf{E}(t) = -\partial_t \mathbf{A}$. A non-obvious consequence of the choice of gauge employed in Eq. 1.12 is that *apparent* unphysical infrared divergences permeate the perturbative results. This can be seen by noting that $\mathbf{A}(\omega) = -i\mathbf{E}(\omega)/\omega$, giving a factor of ω^{-n} in the n-th order nonlinear conductivity (see Section 2.4) that, if not cancelled, leads to an infinite DC response. This happens even for insulators with no free charge carriers present, a clearly unphysical situation. The cancellations do occur, of course, and the divergences are only apparent, as can be shown by careful manipulations. Aspnes [45] provided one of the first demonstrations of this for crystals of zincblende symmetry. The difficulty is there for the nonlinear optical response of atomic systems too [46], but is aggravated in the case of solids by the complexity of the expressions [47]. It sets the perturbation theory on a delicate ground, where approximations easily result in nonsensical answers and it ultimately led this approach to fall in disfavor. This serves also as a warning to the reader: some care should be taken when revisiting this older literature, since these problems were not yet fully understood and would not be appropriately addressed until

[5] Actually, the calculations concerned the second order hyperpolarisability. But for our purposes here, it does not matter whether we are talking about hyperpolarisabilities, susceptibilities or conductivities; it all amounts to the same, since they are easily converted into each other (when ignoring local field effects).

[6] Previous attempts used phenomenological models around the critical points and/or focused only on the optical response at zero frequency.

[7] For the sake of simplicity, theorists are sometimes tempted to take them as constants.

the nineties [47–49], when alternative formulations were proposed and the previous failures were recognized in hindsight as being the result of sum rule violations.

The natural solution was then to adopt a different gauge. Sipe and Ghahramani [47], basing themselves on a previous work [50], did just that and developed an approach that was inherently free from infrared divergences and provided the first generally applicable expressions for the nonlinear conductivity. Alternative approaches were derived from density functional theory, including the works of Levine [51], which introduces a position dependent field and takes the limit to an uniform field at the end of the calculation, and Dal Corso *et al* [52, 53], that opted to perform the calculations in real space with Wannier functions.

However, the most successful and popular method in current use is the one introduced by Aversa and Sipe [49], based on a dipole interaction term:

$$\hat{H} = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}}) - e\,\hat{\mathbf{r}} \cdot \mathbf{E}(t) \tag{1.13}$$

It corresponds to a different, but equivalent, gauge choice than Eq. 1.12. The associated perturbation theory has no spurious divergences, no difficulties emerge from band truncation and it allows for a clean separation of intraband and interband contributions. The representation of the position operator in the Bloch basis is nontrivial [54] (see Section 2.3), but this was proven not to be an obstacle in computing the nonlinear conductivity [48, 49]. With the use of this Hamiltonian and the removal of artificial divergences, actual physical divergences could be identified in the relaxation-free limit that were related to new interesting physical effects, such as the shift and injection currents [5, 49, 55].

As mentioned before, the discussion presented here is confined to systems of non-interacting electrons and spatially uniform electric fields (at the atomic scale). Exciton effects, quadrupole moments and electron-electron interactions are certainly important, but reside outside the scope of this thesis. If this historical review offers any perspective, it should be that enough difficulties are present already at the single electron level to merit a detailed discussion.

The use of Eq. 1.12 is often referred to as the velocity gauge and Eq. 1.13 as the length gauge. The apparent discrepancies in the derived results despite gauge invariance was a decades-long puzzle that only recently got properly resolved [56]. We shall return to this question (Section 2.4), when listing the advantages and disadvantages for each gauge; most of the current text is dedicated to clarifying aspects of these two perturbation methods and presenting additional refinements. We start by revisiting the standard perturbation theory calculation of the nonlinear response.

Chapter 2

Perturbation theory

In order to perform the expansion of the current in the optical fields (Eq. 1.5), we must recur to quantum mechanical perturbation theory. The general case is first reviewed with relevant notation established for future use. The specific nonlinear response functions of interest are then identified and the role of gauge fixing discussed.

2.1 Density matrix formalism

An ensemble of quantum systems can be described by a density operator ρ evolving in time according to the von Neumann equation:

$$i\,\hbar\,\partial_t\hat{\rho} = \left[\hat{H}, \hat{\rho}\right] \tag{2.1}$$

The statistical average of any observable can be obtained by tracing its product with the density operator, $\mathcal{O} = \langle\hat{\mathcal{O}}\rangle = \mathrm{Tr}(\hat{\mathcal{O}}\,\hat{\rho})$, and therefore all the dynamics of the ensemble is contained in Eq. 2.1. Consider the Hamiltonian

$$\hat{H}(t) = \hat{H}_0 + \hat{\mathcal{V}}(t) \tag{2.2}$$

where the first term on the right hand side is assumed to be well understood and have known eigenstates $|\psi_a\rangle$ with energies ϵ_a, and the second term couples a set of observables $\hat{\mathcal{O}}^\alpha$ to external classical fields E^α,

$$\hat{\mathcal{V}}(t) = \hat{\mathcal{O}}^\alpha\, E^\alpha(t) \tag{2.3}$$

This term is assumed to be small and a perturbative treatment will ensue.

In the absence of the perturbation, the system is assumed to be in thermal equilibrium and the density matrix is time-independent and diagonal in the eigenbasis of $\hat{H}_0$. The perturbation is switched on at $t = -\infty$. To describe the time evolution of the ensemble, we must solve Eq. 2.1 with $\hat{\rho}(-\infty) = \hat{\rho}_0$.

In the interaction picture, the von Neumann equation takes the form

$$i\,\hbar\,\partial_t\hat{\rho}_I(t) = \left[\hat{\mathcal{V}}_I(t), \hat{\rho}_I(t)\right] \tag{2.4}$$

with

$$\hat{\rho}_I(t) \equiv \hat{U}_0^\dagger(t)\,\hat{\rho}\,\hat{U}_0(t) \qquad \hat{\mathcal{V}}_I(t) \equiv \hat{U}_0^\dagger(t)\hat{\mathcal{V}}(t)U_0(t) = \hat{\mathcal{O}}_I^\alpha(t)\,E^\alpha(t) \tag{2.5}$$

where $\hat{U}_0(t) \equiv e^{-\frac{i\hat{H}_0}{\hbar}t}$ is the time evolution operator associated to the unperturbed Hamiltonian.

The solution to Eq. 2.4 can be written as

$$\hat{\rho}_I(t) = \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} [\hat{\mathcal{V}}_I(t'), \hat{\rho}_I(t')]\, dt' = \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} [\hat{\mathcal{O}}_I^{\alpha}(t'), \hat{\rho}_I(t')]\, E^{\alpha}(t')\, dt' \tag{2.6}$$

Transforming back to the Schrödinger picture,

$$\begin{aligned}
\hat{\rho}(t) &= \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} \hat{U}_0(t)[\hat{\mathcal{O}}_I^{\alpha}(t'), \hat{\rho}_I(t')]\, \hat{U}_0^{\dagger}(t)\, E^{\alpha}(t')\, dt' \\
&= \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} \hat{U}_0^{\dagger}(-t)[\hat{\mathcal{O}}_I^{\alpha}(t'), \hat{\rho}_I(t')]\, \hat{U}_0(-t)\, E^{\alpha}(t')\, dt' \\
&= \hat{\rho}_0 - \frac{i}{\hbar}\int_{-\infty}^{t} [\hat{\mathcal{O}}_I^{\alpha}(t'-t), \hat{\rho}_I(t'-t)]\, E^{\alpha}(t')\, dt'
\end{aligned} \tag{2.7}$$

From here the usual perturbative approach is followed. In zeroth order in the perturbation, the density operator is described by the equilibrium distribution $\hat{\rho}^{(0)}(t) = \hat{\rho}_0$.

In first order,

$$\hat{\rho}^{(1)}(t) = -\frac{i}{\hbar}\int_{-\infty}^{t} dt'\, [\hat{\mathcal{O}}_I^{\alpha}(t'-t), \hat{\rho}_0]\, E^{\alpha}(t') \tag{2.8}$$

In second order,

$$\hat{\rho}^{(2)}(t) = \left(-\frac{i}{\hbar}\right)^2 \int_{-\infty}^{t} dt_2 \int_{-\infty}^{t_2} dt_1\, [\hat{\mathcal{O}}_I^{\alpha_2}(t_2 - t), [\hat{\mathcal{O}}_I^{\alpha_1}(t_1 - t), \hat{\rho}_0]]\, E^{\alpha_1}(t_1)\, E^{\alpha_2}(t_2) \tag{2.9}$$

Generalizing to order n,

$$\hat{\rho}^{(n)}(t) = \left(-\frac{i}{\hbar}\right)^n \int_{-\infty}^{t} dt_n ... \int_{-\infty}^{t_2} dt_1 [\hat{\mathcal{O}}_I^{\alpha_n}(t_n - t), ...[\hat{\mathcal{O}}_I^{\alpha_1}(t_1 - t), \hat{\rho}_0]...]E^{\alpha_1}(t_1)...E^{\alpha_n}(t_n) \tag{2.10}$$

This is the general structure of the perturbative solutions to Eq. 2.1.

A consequence of the perturbative solutions having the form of nested commutators is that their trace must always be zero. Indeed, it follows from Eq. 2.7 that $\mathrm{Tr}\,\hat{\rho} = \mathrm{Tr}\,\hat{\rho}_0$ and $\mathrm{Tr}\,\hat{\rho}^{(n)} = 0$ for $n > 0$.

These solutions were formulated in the time domain. For the case of an harmonic perturbation or, more generally, in the case that the classical fields have a spectrum localized around some central frequency, it is useful to express the density operator and the perturbation theory in the frequency domain,

$$\hat{\rho}(\omega) \equiv \int_{-\infty}^{+\infty} dt\, \hat{\rho}(t)\, e^{i\omega t} \tag{2.11}$$

The perturbative solutions could be expressed in the frequency domain by Fourier transformation of Eqs. 2.8-2.10, but an easier derivation results by returning to the

equation of motion (Eq. 2.1) and using the representation provided by the eigenstates of H_0:

$$i\,\hbar\,\partial_t \rho_{ab} = [H_0, \rho]_{ab} + [\mathcal{V}, \rho]_{ab} = \rho_{ab}\,\Delta\epsilon_{ab} + [\mathcal{V}, \rho]_{ab} \tag{2.12}$$

with $\Delta\epsilon_{ab} \equiv \epsilon_a - \epsilon_b$. Applying a Fourier transform,

$$(\hbar\omega - \Delta\epsilon_{ab})\ \rho_{ab}(\omega) = \int_{-\infty}^{+\infty} [\mathcal{V}(t), \rho(t)]_{ab} e^{i\omega t}\, dt \tag{2.13}$$

giving the recursion relation

$$\begin{aligned}\rho^{(n)}_{ab}(\omega) &= \frac{1}{\hbar\omega - \Delta\epsilon_{ab}} \int_{-\infty}^{+\infty} [\mathcal{O}^\alpha, \rho^{(n-1)}(t)]_{ab}\, E^\alpha(t)\, e^{i\omega t}\, dt \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi}\frac{d\omega''}{2\pi} \frac{[\mathcal{O}^\alpha, \rho^{(n-1)}(\omega')]_{ab}}{\hbar\omega - \Delta\epsilon_{ab}}\, E^\alpha(\omega'')\,(2\pi)\,\delta(\omega - \omega' - \omega'')\end{aligned} \tag{2.14}$$

Once again, we note that the perturbation is small and that $\rho^{(0)}(\omega) = \rho_0\,(2\pi)\,\delta(\omega)$. In linear order,

$$\rho^{(1)}_{ab}(\omega) = \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\,\pi} \frac{[\mathcal{O}^\alpha, \rho_0]_{ab}}{\hbar\omega_1 - \Delta\epsilon_{ab}}\, E^\alpha(\omega_1)\,(2\,\pi)\,\delta(\omega - \omega_1) = \frac{[\mathcal{O}^\alpha, \rho_0]_{ab}}{\hbar\omega - \Delta\epsilon_{ab}}\, E^\alpha(\omega) \tag{2.15}$$

In second order,

$$\begin{aligned}\rho^{(2)}_{ab}(\omega) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi}\frac{d\omega_2}{2\pi} \frac{[\mathcal{O}^{\alpha_2}, \rho^{(1)}(\omega_1)]_{ab}}{\hbar\omega - \Delta\epsilon_{ab}}\, E^{\alpha_2}(\omega_2)\,(2\pi)\,\delta(\omega - \omega_1 - \omega_2) \\ &= \int_{-\infty}^{+\infty} \frac{d\omega_2}{2\,\pi} \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\,\pi} \frac{1}{\hbar\omega_1 + \hbar\omega_2 - \Delta\epsilon_{ab}} \left[\mathcal{O}^{\alpha_2}, \frac{1}{\hbar\omega_1 - \Delta\epsilon} \circ [\mathcal{O}^{\alpha_1}, \rho_0]\right]_{ab} \\ &\qquad \times E^{\alpha_1}(\omega_1) E^{\alpha_2}(\omega_2)\,(2\pi)\,\delta(\omega - \omega_1 - \omega_2)\end{aligned} \tag{2.16}$$

where $\circ$ stands for the Hadamard product: $(A \circ B)_{ab} = A_{ab}B_{ab}$.

The notation just used is probably unfamiliar to the reader. Since it is often employed in this text, it is here exemplified by making the commutator structure at second order explicit,

$$\left[\mathcal{O}^{\alpha_2}, \frac{1}{\hbar\omega_1 - \Delta\epsilon} \circ [\mathcal{O}^{\alpha_1}, \rho_0]\right]_{ab} = \sum_c \mathcal{O}^{\alpha_2}_{ac} \frac{[\mathcal{O}^{\alpha_1}, \rho_0]_{cb}}{\hbar\omega_1 - \Delta\epsilon_{cb}} - \sum_c \frac{[\mathcal{O}^{\alpha_1}, \rho_0]_{ac}}{\hbar\omega_1 - \Delta\epsilon_{ac}} \mathcal{O}^{\alpha_2}_{cb} \tag{2.17}$$

where, at this point, the inner commutator could also be expanded. This notation is the one used in [56, 57] and is useful in abbreviating expressions.

In general,

$$\begin{aligned}\rho^{(n)}_{ab}(\omega) = \int_{-\infty}^{+\infty} \frac{d\omega_n}{2\,\pi} \ldots \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\,\pi} \frac{1}{\hbar\omega_1 + \ldots + \hbar\omega_n - \Delta\epsilon} \circ \left[\mathcal{O}^{\alpha_n}, \ldots \frac{1}{\hbar\omega_1 - \Delta\epsilon} \circ [\mathcal{O}^{\alpha_1}, \rho_0] \ldots\right]_{ab} \\ \times\ E^{\alpha_1}(\omega_1) \ldots E^{\alpha_n}(\omega_n)\,(2\pi)\,\delta(\omega - \omega_1 \cdots - \omega_n)\end{aligned} \tag{2.18}$$

Often, when performing this type of computations, the commutators are expanded in all their glorious detail, with final expressions containing many terms

and patterns that hard to recognize if not for diagrammatic techniques, when they exist. While the commutators will inevitably be expanded to perform specific computations in any given system, there is great value in maintaining this condensed notation until such is required. General properties are often more easily found and proven by manipulation of these expressions. Much of this thesis work builds on a proper appreciation of the structure of nested commutators in Eq. 2.18 and its generalization.

It is also worth noting that while Eqs. 2.8-2.10 are written in operator form, Eqs. 2.15-2.18 involve matrices in the representation set by the eigenstates of H_0. Nothing was assumed about the spectrum of H_0, which may therefore involve either discrete energy levels or a continuum (bands) or both.

2.2 Response functions

A particular application of perturbation theory is the calculation of response functions (let us call them, say, σ), that describe how some observables of interest $J^\beta \equiv \langle \hat{J}^\beta \rangle = \mathrm{Tr}\Big(\hat{J}^\beta\,\hat{\rho}\Big)$ are affected by the presence of the external fields $E^\alpha(t)$. In nonlinear optics, J would be the electric current and E the electric field, but for now the setting remains completely general.

A notation is introduced at this point that shall prove useful in abbreviating calculations,

$$\hat{\rho}^{(n)}(t) \equiv \int_{-\infty}^{+\infty} dt_n \cdots \int_{-\infty}^{+\infty} dt_1\, \hat{\rho}^{\alpha_1\ldots\alpha_n}(t-t_1,\ldots,t-t_n)\, E^{\alpha_1}(t_1)\ldots E^{\alpha_n}(t_n) \quad (2.19)$$

for $n \geq 1$.

Comparing with Eq. 2.10, we see that

$$\hat{\rho}^{\alpha_1\ldots\alpha_n}(t_1,...,t_n) = \left(-\frac{i}{\hbar}\right)^n [\hat{\mathcal{O}}_I^{\alpha_n}(-t_n),...[\hat{\mathcal{O}}_I^{\alpha_1}(-t_1),\hat{\rho}_0]...]\,\Theta(t_1-t_2)...\Theta(t_{n-1}-t_n)\Theta(t_n) \quad (2.20)$$

Likewise, we can define in the frequency domain,

$$\hat{\rho}^{(n)}(\omega) \equiv \int_{-\infty}^{+\infty} \frac{d\omega_n}{2\,\pi}... \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\,\pi}\, \hat{\rho}^{\alpha_1\ldots\alpha_n}(\omega_1,...,\omega_n)\, E^{\alpha_1}(\omega_1)...E^{\alpha_n}(\omega_n)\,(2\pi)\,\delta(\omega-\omega_1...-\omega_n) \quad (2.21)$$

which, by comparison with Eq. 2.18, gives

$$\rho_{ab}^{\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n) = \frac{1}{\hbar\omega_1+\cdots+\hbar\omega_n-\Delta\epsilon} \circ \left[\mathcal{O}^{\alpha_n},\ldots \frac{1}{\hbar\omega_1-\Delta\epsilon} \circ [\mathcal{O}^{\alpha_1},\rho_0]\ldots\right]_{ab} \quad (2.22)$$

It is now straightforward to express a general response function using the previous definitions. If we consider the relation between a set of observables J^β and the external fields, then for sufficiently weak fields a power series expansion may be appropriate,

$$\langle \hat{J}^{\beta} \rangle(E) = \langle \hat{J}^{\beta} \rangle^{(0)} + \langle \hat{J}^{\beta} \rangle^{(1)} + \langle \hat{J}^{\beta} \rangle^{(2)} + \dots \tag{2.23}$$

where

$$\langle \hat{J}^{\beta} \rangle^{(n)}(t) \equiv \int_{-\infty}^{+\infty} dt_n \cdots \int_{-\infty}^{+\infty} dt_1 \, \sigma^{\beta\alpha_1\dots\alpha_n}(t-t_1,\dots,t-t_n)\, E^{\alpha_1}(t_1)\dots E^{\alpha_n}(t_n) \tag{2.24}$$

It can also be expressed in the frequency domain,

$$\langle \hat{J}^{\beta} \rangle^{(n)}(\omega) \equiv \int_{-\infty}^{+\infty} \frac{d\omega_n}{2\,\pi} \cdots \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\,\pi} \, \sigma^{\beta\alpha_1\dots\alpha_n}(\omega_1,\dots,\omega_n) \times E^{\alpha_1}(\omega_1)\dots E^{\alpha_n}(\omega_n)\,(2\pi)\,\delta(\omega-\omega_1\cdots-\omega_n) \tag{2.25}$$

The relation $\mathbf{J}(\mathbf{E})$ is, in the regime where perturbation theory is valid, captured by the time-domain or frequency domain response functions defined in Eqs. 2.24 and 2.25, respectively. Relating this to previous definitions, it is clear that any such response function will always be obtained from Eqs. 2.20 and 2.22 by tracing over the observable of interest,

$$\sigma^{\beta\alpha_1\dots\alpha_n}(t_1,\dots,t_n) = \mathrm{Tr}(\hat{J}^{\beta}\,\hat{\rho}^{\alpha_1\dots\alpha_n}(t_1,\dots,t_n)) \tag{2.26}$$

$$\sigma^{\beta\alpha_1\dots\alpha_n}(\omega_1,\dots,\omega_n) = \mathrm{Tr}(\hat{J}^{\beta}\,\hat{\rho}^{\alpha_1\dots\alpha_n}(\omega_1,\dots,\omega_n)) \tag{2.27}$$

If written explicitly,

$$\sigma^{\beta\alpha_1\dots\alpha_n}(t_1,...,t_n) = \left(-\frac{i}{\hbar}\right)^n \mathrm{Tr}\left(\hat{J}^{\beta}\,[\hat{\mathcal{O}}_I^{\alpha_n}(-t_n),...[\hat{\mathcal{O}}_I^{\alpha_1}(-t_1),\hat{\rho}_0]...]\right)\,\Theta(t_1-t_2)...\Theta(t_{n-1}-t_n)\Theta(t_n) \tag{2.28}$$

$$\sigma^{\beta\alpha_1\dots\alpha_n}(\omega_1,...,\omega_n) = \sum_{a,b} \frac{J^{\beta}_{ba}}{\hbar\omega_1+...+\hbar\omega_n-\Delta\epsilon_{ab}} \left[\mathcal{O}^{\alpha_n},...\frac{1}{\hbar\omega_1-\Delta\epsilon}\circ[\mathcal{O}^{\alpha_1},\rho_0]\,...\right]_{ab} \tag{2.29}$$

The quantities defined in Eqs. 2.19 and 2.21 are themselves response functions of sort, describing the relation between the density operator (an observable) and the classical fields. But the expressions in Eqs 2.28 and 2.29 are, of course, the ones of greater interest for us here. As highlighted in the previous section, the frequency components of the nonlinear conductivity are particularly useful in nonlinear optics and Eq. 2.29 is the main reason we undertook this review of density matrix perturbation theory.

As a side note, it is useful to know that, when calculating response functions, the structure of the perturbative solutions in Eqs. 2.28 and 2.29 can be rearranged in the following way:

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n) = \left(\frac{i}{\hbar}\right)^n \mathrm{Tr}\left(\hat{\rho}_0\left[\hat{\mathcal{O}}_I^{\alpha_1}(-t_1),\ldots\left[\hat{\mathcal{O}}_I^{\alpha_n}(-t_n),\hat{J}^\beta\right]\ldots\right]\right)\Theta(t_1-t_2)\ldots\Theta(t_{n-1}-t_n)\Theta(t_n) \tag{2.30}$$

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n) = (-1)^n\sum_a(\rho_0)_{aa}\left[\mathcal{O}^{\alpha_1},\frac{1}{\hbar\omega_1+\Delta\epsilon}\circ\ldots\left[\mathcal{O}^{\alpha_n},\frac{1}{\hbar\omega_1+\cdots+\hbar\omega_n+\Delta\epsilon}\circ J^\beta\right]\ldots\right]_{aa} \tag{2.31}$$

Eqs. 2.30 and 2.31 are completely equivalent to Eqs. 2.28 and 2.29. To go from one formulation to another is a matter of changing indices and moving commutators around with the cyclic property of the trace: $\mathrm{Tr}(A[B,C]) = \mathrm{Tr}(C[A,B])$. The perturbation is no longer being commuted with the thermal equilibrium distribution ρ_0, but instead acts on the observable J^β, which can be advantageous. Another advantage of this way of writing the perturbation theory is the decomposition of the response functions into a sum of contributions over each state, since ρ_0 is diagonal in the basis of eigenstates of H_0. This type of expression is common in perturbative treatments of the nonlinear optics of atomic systems [3].

A fairly important point is left to be discussed, as the application of the previous formulas to derive the nonlinear conductivity of a quantum system is not as direct as it may seem. Inspection of Eq. 2.29 suggests that the frequency ω_1 plays a somewhat different role than ω_2, for example, since it is placed in more denominators, but this is not the case. It follows from Eq. 2.25 that only the symmetrical part of the nonlinear conductivity survives the integration and is therefore physical. By symmetrical part it is meant the part that respects *intrinsic permutation symmetry*:

$$\sigma^{\beta\ldots\alpha_i\ldots\alpha_j\ldots}(\ldots,\omega_i,\ldots,\omega_j,\ldots) = \sigma^{\beta\ldots\alpha_j\ldots\alpha_i\ldots}(\ldots,\omega_j,\ldots,\omega_i,\ldots) \tag{2.32}$$

for any $i,j\in\{1,\ldots,n\}$.

Any function that does not respect intrinsic permutation symmetry could be added to the nonlinear conductivity without affecting the resulting current. When wishing to infer physical information from the nonlinear conductivity, it is therefore an absolutely necessary step to symmetrize the expression in Eq. 2.29 to ensure that the conductivity obeys intrinsic permutation symmetry. In second order, for instance,

$$\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \sigma_S^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1) \tag{2.33}$$

must be guaranteed by extracting the symmetrical part of the conductivities derived from Eq. 2.29 with $n=2$,

$$\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{1}{2}\left(\sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)+\sigma^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1)\right) \tag{2.34}$$

2.3 The covariant derivative

Returning to our problem of finding the nonlinear optical response of crystalline solids, we apply the formalism just developed to the dynamics of an electron moving

in the periodic potential of a crystal:

$$\hat{H}_0 = \frac{\hat{p}^2}{2\,m} + V(\hat{\mathbf{r}}) \tag{2.35}$$

where $\hat{p} = ||\hat{\mathbf{p}}||$ and $V(\hat{\mathbf{r}})$ has the periodicity of a given Bravais lattice $\{\mathbf{R}_n\}$.

Bloch's theorem states that the eigenfunctions of an Hamiltonian $\hat{H}_0$ with discrete translation symmetry have the form

$$\psi_{\mathbf{k}a}(\boldsymbol{r}) = e^{i\,\mathbf{k}\cdot\boldsymbol{r}}\, u_{\mathbf{k}a}(\boldsymbol{r}) \tag{2.36}$$

where $u_{\mathbf{k}a}(\boldsymbol{r})$ is periodic in the Bravais lattice $u_{\mathbf{k}a}(\boldsymbol{r} + \boldsymbol{R}_n) = u_{\mathbf{k}a}(\boldsymbol{r})$, $\mathbf{k}$ is the Bloch vector and a the band index [58]. The Bloch vector $\mathbf{k}$ is a good quantum number that, in the thermodynamic limit, takes any value in the First Brillouin Zone (FBZ). In this limit, sums over $\mathbf{k}$ become integrals over the FBZ. For instance, the closure relation

$$\sum_{\mathbf{k}}\sum_{a} |\psi_{\mathbf{k}a}\rangle\,\langle\psi_{\mathbf{k}a}| = \hat{\mathbb{1}} \to \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a} |\psi_{\mathbf{k}a}\rangle\,\langle\psi_{\mathbf{k}a}| = \hat{\mathbb{1}} \tag{2.37}$$

where d is the dimensionality of the system (and its FBZ).

In the eigenvalue equation,

$$\hat{H}_0\,|\psi_{\mathbf{k}a}\rangle = \epsilon_{\mathbf{k}a}\,|\psi_{\mathbf{k}a}\rangle \tag{2.38}$$

the energies $\epsilon_{\mathbf{k}a}$ are defined continuously over the FBZ, forming multiple energy bands, labeled by the band index a.

With these definitions, the Bloch basis is normalized according to

$$\langle\psi_{\mathbf{k}a}|\psi_{\mathbf{k}'b}\rangle = (2\pi)^d\,\delta_{ab}\,\delta(\mathbf{k}-\mathbf{k}') \tag{2.39}$$

Operators that are invariant under the lattice translations are diagonal in $\mathbf{k}$ and their matrix elements are denoted as follows

$$\langle\psi_{\mathbf{k}a}|\,\hat{\mathcal{O}}\,|\psi_{\mathbf{k}'b}\rangle = \mathcal{O}_{\mathbf{k}ab}\,(2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}') \tag{2.40}$$

As an example, the equilibrium distribution for a system of independent electrons is the Fermi-Dirac distribution

$$\hat{\rho}_0 = f(\hat{H}_0) \equiv \frac{1}{1+e^{(\hat{H}_0-\mu)/k_BT}} \tag{2.41}$$

with μ as the chemical potential, and it is clearly diagonal in the eigenbasis of H_0,

$$\langle\psi_{\mathbf{k}a}|\,\hat{\rho}_0\,|\psi_{\mathbf{k}'b}\rangle = (\rho_0)_{\mathbf{k}ab}\;(2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}') = f_{\mathbf{k}a}\,\delta_{ab}\,(2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}') \tag{2.42}$$

where

$$f_{\mathbf{k}a} \equiv f(\epsilon_{\mathbf{k}a}) = \frac{1}{1+e^{(\epsilon_{\mathbf{k}a}-\mu)/k_BT}} \tag{2.43}$$

The carrier concentration n_e is determined by the chemical potential, with the density matrix normalized so that $\mathrm{Tr}(\hat{\rho}_0) = n_e$.

The full Hamiltonian $\hat{H}$ includes not only the crystalline potential but also a term $\hat{\mathcal{V}}(t)$ that couples the quantum electron motion to the classical optical field. The specific form of this coupling will depend on our choice of gauge. For the length

gauge Hamiltonian in Eq. 1.13, the coupling has the form we have been considering (Eq. 2.3), where E^α is the electric field and the observable $\hat{\mathcal{O}}^\alpha$ corresponds to the electric dipole moment:

$$\hat{\mathcal{V}}(t) = e\,\hat{r}^\alpha\,E^\alpha(t) \tag{2.44}$$

To use perturbation theory and find the nonlinear conductivities of this system, it is necessary to write the perturbation in the eigenbasis of the unperturbed Hamiltonian H_0. That is, it is required to write the position operator in the Bloch basis [54],

$$\begin{aligned}
&\langle\psi_{\mathbf{k}'a}|\,\hat{r}^\alpha\,|\psi_{\mathbf{k}b}\rangle = \int d^3\mathbf{r}\,\psi^*_{\mathbf{k}'a}(\mathbf{r})\,r^\alpha\,\psi_{\mathbf{k}b}(\mathbf{r}) \\
&= \int d^3\mathbf{r}\,e^{-i\,\mathbf{k}'\cdot\mathbf{r}}\,u^*_{\mathbf{k}'a}(\mathbf{r})\,r^\alpha\,e^{i\,\mathbf{k}\cdot\mathbf{r}}\,u_{\mathbf{k}b}(\mathbf{r}) = \int d^3\mathbf{r}\,e^{-i\,\mathbf{k}'\cdot\mathbf{r}}\,u^*_{\mathbf{k}'a}(\mathbf{r})\left(-i\,\frac{\partial}{\partial k_\alpha}\,e^{i\,\mathbf{k}\cdot\mathbf{r}}\right)u_{\mathbf{k}b}(\mathbf{r}) \\
&= -i\,\frac{\partial}{\partial k_\alpha}\int d^3\mathbf{r}\,e^{-i\,\mathbf{k}'\cdot\mathbf{r}}\,u^*_{\mathbf{k}'a}(\mathbf{r})\,e^{i\,\mathbf{k}\cdot\mathbf{r}}\,u_{\mathbf{k}b}(\mathbf{r}) + i\int d^3\mathbf{r}\,e^{-i\,\mathbf{k}'\cdot\mathbf{r}}\,u^*_{\mathbf{k}'a}(\mathbf{r})\,e^{i\,\mathbf{k}\cdot\mathbf{r}}\,\frac{\partial u_{\mathbf{k}b}(\mathbf{r})}{\partial k_\alpha} \\
&= -i\,\frac{\partial}{\partial k_\alpha}\,\langle\psi_{\mathbf{k}'a}|\psi_{\mathbf{k}b}\rangle + i\sum_{\mathbf{R}_n} e^{i\,(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_n}\int_{u.\,c.} d^3\mathbf{r}\,u^*_{\mathbf{k}'a}(\mathbf{r})\,\frac{\partial u_{\mathbf{k}b}(\mathbf{r})}{\partial k_\alpha}
\end{aligned} \tag{2.45}$$

where we used the periodicity of the $u_{\mathbf{k}a}(\mathbf{r})$ functions to separate the integral over all real space into an integral over the unit cell and a sum over the Bravais lattice sites. The lattice sum evaluates to a Dirac delta,

$$\sum_{\mathbf{R}_n} e^{i\,\mathbf{k}\cdot\mathbf{R}_n} = \frac{(2\pi)^d}{v_c}\,\delta(\mathbf{k}) \tag{2.46}$$

where v_c is the volume of the unit cell. Substituting in Eq. 2.45,

$$= -i\,\delta_{ab}\,(2\pi)^d\,\partial^\alpha\delta(\mathbf{k}-\mathbf{k}') + (2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}')\left(\frac{i}{v_c}\int_{u.\,c.} d^3\mathbf{r}\,u^*_{\mathbf{k}a}(\mathbf{r})\,\partial^\alpha u_{\mathbf{k}b}(\mathbf{r})\right) \tag{2.47}$$

Henceforth, the derivative notation is abbreviated: $\partial^\alpha \equiv \partial/\partial k_\alpha$.

The first term in Eq. 2.47 makes it clear that the position operator is not diagonal in $\mathbf{k}$, linking states with neighbouring Bloch vectors. The quantity in parenthesis is closely related to the geometrical phase [59] and is a well-known object in the exploration of geometry in electronic structure theory [60]: the non-abelian Berry connection

$$\mathcal{A}^\alpha_{\mathbf{k}ab} = i\,\langle u_{\mathbf{k}a}|\partial^\alpha u_{\mathbf{k}b}\rangle \equiv \frac{i}{v_c}\int_{u.\,c.} d^3\mathbf{r}\,u^*_{\mathbf{k}a}(\mathbf{r})\,\partial^\alpha u_{\mathbf{k}b}(\mathbf{r}) \tag{2.48}$$

Replacing it in Eq. 2.47,

$$\begin{aligned}
&= -i\,\delta_{ab}\,(2\pi)^d\,\partial^\alpha\delta(\mathbf{k}-\mathbf{k}') + (2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}')\,\mathcal{A}^\alpha_{\mathbf{k}ab} \\
&= (2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}')\,i\,\left(\delta_{ab}\,\partial^\alpha - i\,\mathcal{A}^\alpha_{\mathbf{k}ab}\right) \\
&= (2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}')\,i\,D^\alpha_{\mathbf{k}ab}
\end{aligned} \tag{2.49}$$

where we used integration by parts to write the position operator's matrix elements in terms of the covariant derivative:

$$D^{\alpha}_{\mathbf{k}ab} \equiv \delta_{ab}\,\partial^{\alpha} - i\,\mathcal{A}^{\alpha}_{\mathbf{k}ab} \tag{2.50}$$

It's standard in quantum mechanics to take the momentum operator as the derivative in real space. And vice-versa, the position operator in the eigenbasis of momentum is a derivative in reciprocal space. If instead of plane waves, Bloch waves are considered, the position operator becomes a covariant derivative. In operator form, we summarize the previous derivation with the statement $\hat{r}^{\alpha} = i\,\hat{D}^{\alpha}$.

If some of the steps in the previous derivation look dubious, this is due to the position operator not being strictly defined in the Bloch basis. The equalities are to be understood as identities between distributions and can be made rigorous with the use of test functions.

It should be noted that the difficulty is defining the position operator in the Bloch basis was not entirely unexpected. Under periodic boundary conditions, in a finite crystal, the position operator is notably undefined in the space generated by the Bloch states. The position operator is also not invariant under lattice translations, which means that the perturbation to our Hamiltonian, in a length gauge formulation, breaks the translation symmetry of the lattice.

However, all this turns out to be no real obstacle in developing the perturbation theory, since the perturbation, and the position operator, appear only inside commutators and these are well-defined [49, 57], even when the perturbation itself isn't:

$$\begin{aligned}[\hat{D}^{\alpha}, \hat{\mathcal{O}}] &= \hat{\partial}^{\alpha}\,\hat{\mathcal{O}} - \hat{\mathcal{O}}\,\hat{\partial}^{\alpha} - i\,[\hat{\mathcal{A}}^{\alpha}, \hat{\mathcal{O}}] \\ &= (\hat{\partial}^{\alpha}\hat{\mathcal{O}}) - i\,[\hat{\mathcal{A}}^{\alpha}, \hat{\mathcal{O}}]\end{aligned} \tag{2.51}$$

whose matrix elements are well-defined and diagonal in $\mathbf{k}$,

$$[D^{\alpha}, \mathcal{O}]_{\mathbf{k}ab} = (\partial^{\alpha}\mathcal{O}_{\mathbf{k}ab}) - i\sum_{c}\left(\mathcal{A}^{\alpha}_{\mathbf{k}ac}\mathcal{O}_{\mathbf{k}cb} - \mathcal{A}^{\alpha}_{\mathbf{k}cb}\mathcal{O}_{\mathbf{k}ac}\right) \tag{2.52}$$

An useful example is the covariant differentiation of the equilibrium distribution,

$$[D^{\alpha}, \rho_0]_{\mathbf{k}ab} = \partial^{\alpha} f_{\mathbf{k}a}\,\delta_{ab} - i\,\mathcal{A}^{\alpha}_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba} \tag{2.53}$$

with $\Delta f_{\mathbf{k}ba} \equiv f_{\mathbf{k}b} - f_{\mathbf{k}a}$.

Another instance of a commutator with the position operator comes from evaluating the matrix elements of the current operator,

$$\hat{J}^{\beta} = -e\,\hat{v}^{\beta} = \frac{i\,e}{\hbar}\,[\hat{r}^{\beta}, \hat{H}] = \frac{i\,e}{\hbar}\,[\hat{r}^{\beta}, \hat{H}_0] = -\frac{e}{\hbar}\,[\hat{D}^{\beta}, \hat{H}_0] \tag{2.54}$$

which in the Bloch basis becomes,

$$J^{\beta}_{\mathbf{k}ab} = -\frac{e}{\hbar}\,[D^{\beta}, H_0]_{\mathbf{k}ab} = -\frac{e}{\hbar}\left(\partial^{\beta}\epsilon_{\mathbf{k}a}\,\delta_{ab} + i\,\mathcal{A}^{\beta}_{\mathbf{k}ab}\,\Delta\epsilon_{\mathbf{k}ab}\right) \tag{2.55}$$

with $\Delta\epsilon_{\mathbf{k}ab} \equiv \epsilon_{\mathbf{k}a} - \epsilon_{\mathbf{k}b}$.

The emergence of the covariant derivative is deeply related to the existence of another gauge symmetry, beyond that of the classical electromagnetic field. Any physical prediction of the theory should remain invariant under the transformation

$$|\psi_{\mathbf{k}a}\rangle \to e^{i\,\phi_{\mathbf{k}a}}\,|\psi_{\mathbf{k}a}\rangle \tag{2.56}$$

However, the $\mathbf{k}$-derivative transforms as[1]

$$\langle\psi_{\mathbf{k}a}|\,i\,\hat{\partial}^{\alpha}\,|\psi_{\mathbf{k}b}\rangle \to e^{i(\phi_{\mathbf{k}b}-\phi_{\mathbf{k}a})}\,\langle\psi_{\mathbf{k}a}|\,i\,\hat{\partial}^{\alpha}\,|\psi_{\mathbf{k}b}\rangle + \delta_{ab}\,(\partial^{\alpha}\phi_{\mathbf{k}a})\,\langle\psi_{\mathbf{k}a}|\psi_{\mathbf{k}a}\rangle \tag{2.57}$$

This is compensated by the transformation properties of the Berry connection,

$$\mathcal{A}^{\alpha}_{\mathbf{k}ab} \to e^{-i\phi_{\mathbf{k}a}}\,\mathcal{A}^{\alpha}_{\mathbf{k}ab}\,e^{i\phi_{\mathbf{k}b}} - \delta_{ab}\,(\partial^{\alpha}\phi_{\mathbf{k}a}) \tag{2.58}$$

implying that the covariant derivative will transform as

$$D^{\alpha}_{\mathbf{k}ab} \to e^{i(\phi_{\mathbf{k}b}-\phi_{\mathbf{k}a})}\,D^{\alpha}_{\mathbf{k}ab} \tag{2.59}$$

leaving only the phase factors. These cancel whenever traces are evaluated (Eq. 2.29).

In this way, gauge invariance of physical predictions under these phase transformations is ensured.

2.4 A tale of two gauges

2.4.1 Length gauge

Once the length gauge perturbation (Eq. 2.44) is written in the Bloch basis, the previous formulas (Eq. 2.29) for nonlinear response functions can be used, immediately giving the linear and nonlinear conductivities upon substitution of the observable $\mathcal{O}^{\alpha}$ by the dipole operator $e\,r^{\alpha} = i\,e\,D^{\alpha}$. The resulting formula for the optical conductivity is equivalent to Kubo's formula [61], when expressed in frequency space,

$$\begin{aligned}\sigma^{\beta\alpha}(\omega) &= \mathrm{Tr}\left(\hat{J}^{\beta}\,\hat{\rho}^{\alpha}(\omega)\right)\\ &= e\int\frac{d^{d}\mathbf{k}}{(2\pi)^{d}}\sum_{a,b}\frac{J^{\beta}_{\mathbf{k}ba}\,[r^{\alpha},\rho_{0}]_{\mathbf{k}ab}}{\hbar\omega-\Delta\epsilon_{\mathbf{k}ab}}\\ &= i\,e\int\frac{d^{d}\mathbf{k}}{(2\pi)^{d}}\sum_{a,b}\frac{J^{\beta}_{\mathbf{k}ba}\,[D^{\alpha},\rho_{0}]_{\mathbf{k}ab}}{\hbar\omega-\Delta\epsilon_{\mathbf{k}ab}}\end{aligned} \tag{2.60}$$

Replacing the current for the commutator in Eq. 2.54,

$$\sigma^{\beta\alpha}(\omega) = -\frac{i\,e^{2}}{\hbar}\int\frac{d^{d}\mathbf{k}}{(2\pi)^{d}}\sum_{a,b}\frac{\left[D^{\beta},H_{0}\right]_{\mathbf{k}ba}}{\hbar\omega-\Delta\epsilon_{\mathbf{k}ab}}\,[D^{\alpha},\rho_{0}]_{\mathbf{k}ab} \tag{2.61}$$

[1] Remember that the "matrix elements" of the derivative actually represent the operation of differentiating everything to its right:

$$\langle\psi_{\mathbf{k}a}|\,i\,\hat{\partial}^{\alpha}\,|\psi\rangle = \int\frac{d^{d}\mathbf{k}'}{(2\pi)^{d}}\,\langle\psi_{\mathbf{k}a}|\,i\,\hat{\partial}^{\alpha}\,|\psi_{\mathbf{k}'b}\rangle\,\langle\psi_{\mathbf{k}'b}|\psi\rangle = i(\partial^{\alpha}\,\langle\psi_{\mathbf{k}a}|\psi\rangle) \to (i(\partial^{\alpha}\,\langle\psi_{\mathbf{k}a}|\psi\rangle) + (\partial^{\alpha}\phi_{\mathbf{k}a})\,\langle\psi_{\mathbf{k}a}|\psi\rangle)\,e^{-i\,\phi_{\mathbf{k}a}}$$

This is a fairly compact expression for the optical conductivity. As far as recognizing patterns and building general proofs, this type of expression can be quite useful. But if one desires to compute the nonlinear optical response of a specific material, the commutators must be expanded (Eqs. 2.53 and 2.55) and the dependence on the band energies and Berry connections exposed in detail. Since it is instructive to inspect the formula for the conductivity when written explicitly (and this is easy to do in linear order):

$$\sigma^{\beta\alpha}(\omega) = -\frac{i\,e^2}{\hbar^2\,\omega}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a \partial^\beta\epsilon_{\mathbf{k}a}\,\partial^\alpha f_{\mathbf{k}a} + \frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathcal{A}^\beta_{\mathbf{k}ba}\mathcal{A}^\alpha_{\mathbf{k}ab}\Delta\epsilon_{\mathbf{k}ab}}{\hbar\omega-\Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{2.62}$$

The first and second terms in Eq. 2.62 correspond to intraband and interband transitions, respectively. In a system with completely filled/empty bands (insulator), no intraband transitions are possible and only the second term survives. On the other hand, the linear response of a metallic or semiconductor system with free carriers is dominated by the first term at low frequencies ($\hbar\omega$ much smaller than the gap). In the length gauge, it is always possible to discern the electron motion in terms of a succession of intra- and interband transitions [49], but it becomes increasingly less transparent in the final expressions as higher orders of perturbation theory are considered.

The second order conductivity is obtained from Eq. 2.29 for $n = 2$,

$$\sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{e^3}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\big[D^\beta,H_0\big]_{\mathbf{k}ba}}{\hbar\omega_1+\hbar\omega_2-\Delta\epsilon_{\mathbf{k}ab}}\left[D^{\alpha_2},\frac{1}{\hbar\omega_1-\Delta\epsilon}\circ\big[D^{\alpha_1},\rho_0\big]\right]_{\mathbf{k}ab} \tag{2.63}$$

A detailed treatment of this expression can be found in Appendix F. Right now, it is enough to note some differences with the linear case that already indicate a significant increase in complexity. The first, unavoidable complication is that we are dealing with a higher rank tensor, with more elements to consider. Second, in addition to a purely intraband contribution (where only the derivative part of the position operator is applied) and a purely interband contribution (where only the Berry connection part of the position operator is used), there are mixed terms related to processes where both intra and interband transitions occur. For instance, an electron might first transition from a valence to a conduction band and then have an intraband transition to a neighbouring $\mathbf{k}$-point. This leads to the appearance of derivatives of Berry connections[2] and other complications. Third, when moving past linear order, the derivatives will act on the energy denominators as well: notice how ∂^{α_2} will act on $(\hbar\omega_1-\Delta\epsilon)^{-1}$ in Eq. 2.63. This will enable higher order poles to emerge, further complicating matters. None of these problems are beyond our ability to manage, but they are what makes a computation of the nonlinear conductivity nontrivial.

Finally, we consider the third order conductivity,

[2] More precisely, generalized derivatives of Berry connections [49] (see Chapter 5).

$$\sigma^{\beta\alpha_1\alpha_2\alpha_3}(\omega_1,\omega_2,\omega_3) =$$
$$\frac{i\,e^4}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\left[D^\beta,H_0\right]_{\mathbf{k}ba}}{\hbar\omega_1+\hbar\omega_2+\hbar\omega_3-\Delta\epsilon_{\mathbf{k}ab}}\left[D^{\alpha_3},\frac{1}{\hbar\omega_1+\hbar\omega_2-\Delta\epsilon}\circ\left[D^{\alpha_2},\frac{1}{\hbar\omega_1-\Delta\epsilon}\circ\left[D^{\alpha_1},\rho_0\right]\right]\right]_{\mathbf{k}ab} \tag{2.64}$$

And the pattern is clear for anyone wishing to pursue higher orders of perturbation theory.

In principle, it seems this would already conclude one of the central topics of this work: how to efficiently compute the nonlinear optical conductivity of crystals? Eqs. 2.63 and 2.64 do provide appropriate formulae for such endeavour, but the condensed notation hides a lot of complexity.

If $V(\hat{\mathbf{r}})$ stood for an atomic potential, similar expressions to Eqs. 2.63 and 2.64 would describe the nonlinear optical response for atomic systems. There would be no integration over the FBZ and the position operator would not be a covariant derivative (r_{ab} are then well-defined matrix elements), but all else would remain unchanged. In such case, one could fully expand the commutators and, with a proper knowledge of the energy levels and dipole matrix elements, obtained from *ab initio* calculations or spectroscopic measurements, the dispersion of the nonlinear conductivity could be accurately predicted[3]. Even the complexity of a general n-th order conductivity would not seem too bad, with diagrammatic methods that can identify the various terms and the possibility of using symmetry to generate them from a single term[4].

Unfortunately, the situation is very different for crystals precisely because of the presence of a covariant derivative and the need to perform an integration. Atomic and free electron systems are special cases in a sense, obtained upon taking the appropriate limits. If the derivative part of the position operator is ignored, leaving only the non-abelian Berry connection $\mathcal{A}$, then at each $\mathbf{k}$ in the FBZ the response is identical to that of an atomic system. If only the $\mathbf{k}$-derivative part of the position operator is retained, a purely intraband contribution will be all that is left and the nonlinear response will be akin to that of a free electron gas. The presence of the crystalline potential is then reflected only in the dispersion relation $\epsilon_{\mathbf{k}a}$. This type of nonlinear response results purely from intraband motion and could be directly derived via a Boltzmann equation approach.

In crystals, where the perturbation is the full covariant derivative, the dynamics is more interesting and more complex. Whenever one desires to go a step higher in perturbation theory, the entirety of the expression for the density matrix at previous order must be differentiated, and also commuted with $\mathcal{A}$, implying a very steep increase in the number of terms to evaluate and the complexity of the objects that are to be evaluated. In Chapter 5, we will deal with this complexity and attempt to cast the nonlinear conductivity in as simple a form as possible by breaking it into smaller pieces, according with existing resonances in the FBZ. For the moment, the inherent difficulties of the length gauge should provide sufficient justification to consider other possibilities.

[3] Assuming the independent electron and electric dipole approximations are reasonable.
[4] See pages 70-74 (Section 4.3) of [3].

2.4.2 Velocity gauge

Equivalent but distinct descriptions of the dynamics of a quantum system can be reached by performing unitary transformations on the states and operators: $|\psi\rangle \to \hat{\mathcal{U}}|\psi\rangle$ and $\hat{\mathcal{O}} \to \hat{\mathcal{U}}\hat{\mathcal{O}}\hat{\mathcal{U}}^\dagger$. The unitary transformation may be time-dependent, in which case the Hamiltonian transforms as follows (see Appendix A)

$$\hat{H} \to \hat{\mathcal{U}}(t)\hat{H}\hat{\mathcal{U}}^\dagger(t) + i\hbar(\partial_t\hat{\mathcal{U}}(t))\hat{\mathcal{U}}^\dagger(t) \tag{2.65}$$

A particular time-dependent unitary transformation $\hat{\mathcal{U}}(t) = e^{-ie\hat{r}^\alpha A^\alpha(t)/\hbar}$, with $-\partial_t A(t) = E(t)$, applied to Eq. 1.13 leads to the familiar minimal coupling Hamiltonian (Eq. 1.12),

$$\hat{H} = \frac{(\hat{\mathbf{p}} + e\boldsymbol{A}(t))^2}{2m} + \hat{V}(\hat{\mathbf{r}}) = \frac{\hat{p}^2}{2m} + \hat{V}(\hat{\mathbf{r}}) + \frac{e\hat{p}^\alpha A^\alpha(t)}{m} + \frac{e^2A(t)^2}{2m} \tag{2.66}$$

where we identify the perturbation

$$\hat{\mathcal{V}}(t) = \frac{e}{m}\hat{p}^\alpha A^\alpha(t) + \frac{e^2A(t)^2}{2m} \tag{2.67}$$

The last piece in Eq. 2.67 is neglected as it merely introduces a time-dependent phase shift in the wave functions and has no impact on the dynamics of the density matrix or the other observables of the theory[5]. This form of the perturbation is alluded to in the literature as the "velocity gauge", presumably because the coupling in Eq. 2.67 is close[6] to $e\,v^\alpha(t)\,A^\alpha(t)$. The passage from the length to the velocity gauge is described in greater detail in Appendix A.

Once again, the coupling has the form in Eq. 2.3 with $\hat{\mathcal{O}}^\alpha = e\hat{p}^\alpha/m$ and A^α as the classical field. This seems to be a much simpler perturbation than the length gauge version, with well defined matrix elements in the Bloch basis. Indeed, the momentum operator could be expressed using the following commutator

$$\hat{p}^\alpha = -\frac{im}{\hbar}[\hat{r}^\alpha, \hat{H}_0] \tag{2.68}$$

which is diagonal in $\mathbf{k}$. The commutator may be expanded,

$$\begin{aligned} p^\alpha_{\mathbf{k}ab} &= -\frac{im}{\hbar}[r^\alpha, H_0]_{\mathbf{k}ab} \\ &= \frac{m}{\hbar}[D^\alpha, H_0]_{\mathbf{k}ab} \\ &= \frac{m}{\hbar}\left(\partial^\alpha\epsilon_{\mathbf{k}a}\,\delta_{ab} + i\,\mathcal{A}^\alpha_{\mathbf{k}ab}\,\Delta\epsilon_{\mathbf{k}ab}\right) \end{aligned} \tag{2.69}$$

thereby reducing the ingredients necessary to an evaluation of the nonlinear conductivity down to a knowledge of the energies and the non-abelian Berry connection, same as in the length gauge.

At this point, it seems the substitution of $\hat{\mathcal{O}}^\alpha$ by $e\hat{p}^\alpha/m$ in Eq. 2.29 would return expressions for the nonlinear conductivity. However, the velocity gauge formulation deviates from the perturbative treatment of Section 2.2 in a few ways. A minor point

[5] An easy way to discern this is to note that in any equation of motion the Hamiltonian always appears inside a commutator. The term in question, being a number, commutes with any operator.

[6] The difference is just an number, since $v(t) = (p + e\,A(t))/m$. See Eq. 2.70

is that it is written in terms of the vector potential $\mathbf{A}(t)$ and not the electric field as intended, but this is no obstacle, since in the frequency domain $\mathbf{A}(\omega) = -i\,\mathbf{E}(\omega)/\omega$. As such, it should be enough to include the factor $(-i)^n/\omega_1 \ldots \omega_n$ in the n-the order nonlinear conductivity to account for the conversion of the potential vector into an electric field. A more fundamental difference of the velocity gauge approach lies in the current (or velocity) operator that has now a direct dependence on the optical field:

$$\hat{J}^\beta(t) = -e\,\hat{v}^\beta = \frac{i\,e}{\hbar}\,[\hat{r}^\beta, \hat{H}] = \frac{i\,e}{\hbar}\,[\hat{r}^\beta, \hat{H}_0] - \frac{e^2}{m}\,A^\beta(t) \tag{2.70}$$

Chapter 4 will expand on this point, but here this can be taken into account with only a slight change in the linear response function formula. The ensemble average of the current is

$$\begin{aligned} J^\beta(t) &= \mathrm{Tr}\left(\hat{J}^\beta\,\hat{\rho}\right) = \frac{i\,e}{\hbar}\,\mathrm{Tr}\left([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}\right) - \frac{e^2}{m}\,\mathrm{Tr}(\hat{\rho})\,A^\beta(t) \\ &= \frac{i\,e}{\hbar}\,\mathrm{Tr}\left([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}\right) - \frac{e^2}{m}\,\mathrm{Tr}(\hat{\rho}_0)\,A^\beta(t) \\ &= \frac{i\,e}{\hbar}\,\mathrm{Tr}\left([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}\right) - \frac{n_e\,e^2}{m}\,A^\beta(t) \end{aligned} \tag{2.71}$$

The last term in Eq. 2.71 is linear in the potential vector and therefore contributes to perturbation theory only in first order. As a result, there is an additional term in the linear conductivity,

$$\sigma^{\beta\alpha}(\omega) = \frac{e^2}{m\,\hbar\,\omega}\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\left[r^\beta, H_0\right]_{\mathbf{k}ba}}{\hbar\omega - \Delta\epsilon_{\mathbf{k}ab}}\,\left[p^\alpha, \rho_0\right]_{\mathbf{k}ab} + \frac{i\,n_e\,e^2}{m\,\omega}\,\delta^{\beta\alpha} \tag{2.72}$$

where δ is the Kronecker delta.

This is the velocity gauge version of Eq. 2.61. A curious aspect of this way of writing it is that the Drude conductivity promptly follows when interband transitions are ignored. Pulling from the results of the next subsection, we introduce relaxation in the previous formula by making the substitution $\omega \to \omega + i/\tau$ with $\tau = \gamma^{-1} > 0$. If we additionally set the off-diagonal momentum matrix elements to zero (decoupling the electronic bands),

$$\sigma^{\beta\alpha}(\omega) = \frac{i\,n_e\,e^2}{m\,(\omega + i/\tau)}\,\delta^{\beta\alpha} = \frac{n_e\,e^2\,\tau}{m\,(1 - i\,\omega\,\tau)}\,\delta^{\beta\alpha} = \frac{\sigma^{\beta\alpha}(0)}{1 - i\,\omega\,\tau}\,\delta^{\beta\alpha} \tag{2.73}$$

a known result from Drude's theory of metals.

The derivation of the second order conductivity is already parallel to the length gauge formulation,

$$\begin{aligned} &\sigma^{\beta\alpha_1\alpha_2}(\omega_1, \omega_2) = \\ &-\frac{i\,e^3}{\hbar\,m^2\,\omega_1\,\omega_2}\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\left[r^\beta, H_0\right]_{\mathbf{k}ba}}{\hbar\omega_1 + \hbar\omega_2 - \Delta\epsilon_{\mathbf{k}ab}}\,\left[p^{\alpha_2}, \frac{1}{\hbar\omega_1 - \Delta\epsilon} \circ \left[p^{\alpha_1}, \rho_0\right]\right]_{\mathbf{k}ab} \end{aligned} \tag{2.74}$$

and the same is true for third order,

$$\sigma^{\beta\alpha_1\alpha_2\alpha_3}(\omega_1,\omega_2,\omega_3) = -\frac{e^4}{\hbar\, m^3\, \omega_1\, \omega_2\, \omega_3} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\left[r^\beta, H_0\right]_{\mathbf{k}ba}}{\hbar\omega_1 + \hbar\omega_2 + \hbar\omega_3 - \Delta\epsilon_{\mathbf{k}ab}} \times \left[p^{\alpha_3}, \frac{1}{\hbar\omega_1 + \hbar\omega_2 - \Delta\epsilon} \circ \left[p^{\alpha_2}, \frac{1}{\hbar\omega_1 - \Delta\epsilon} \circ \left[p^{\alpha_1}, \rho_0\right]\right]\right]_{\mathbf{k}ab} \tag{2.75}$$

There are no derivatives in Eqs. 2.72, 2.74 and 2.75, except for those required in assessing the momentum matrix elements (Eq. 2.69). The nonlinear conductivity is a sum over independent contributions, one for each **k**-point in the FBZ. Hence, the nonlinear optical response, when expressed in the velocity gauge, is identical to that of a collection of atoms, each labeled by a Bloch vector. This is no surprise, since the minimal coupling Hamiltonian preserves the translation symmetry of the lattice and the Bloch vector **k** remains a good quantum number even after the coupling with the optical field is introduced.

2.4.3 Length vs velocity gauge

What effectively happens when moving from a length to a velocity gauge description is that the unitary transformation $\mathcal{U}(t)$ decouples the system in **k**. Because the transformation is unitary, the two descriptions are nonetheless entirely equivalent. This can be further confirmed by manipulating the nonlinear conductivity expressions from one formulation to another, showing that they are, in fact, the *same expressions*, just written rather differently (e. g. with some cycling of commutators, relabelling of indices and, most importantly, the use of a few sum rules [57], Eq. 2.75 can be shown to be one and the same as Eq. 2.64).

Unfortunately, the fundamental principle of gauge invariance was not enough to prevent published results using length and velocity gauge methods to differ substantially. This was a source of confusion and debate for a while, but eventually it became clear that the difficulty resided in the approximations that are inevitably made in any calculation of the optical properties of matter.

The Hilbert space of the Hamiltonian in Eq. 2.35 is infinite dimensional. A computation of the optical response then requires an infinite number of states, for atomic and condensed matter systems alike. This is hardly practical. In practice, one always limits the analysis to a subspace of the space of states that is assumed to properly describe the physics. In the case of crystals, the valence and conduction bands nearest to the Fermi level are expected to be sufficient for any computation of physical properties. But the validity of such approximations is seldom ensured.

The equivalence of the length and velocity gauge is broken by band truncation. Perhaps the simplest way to expose this is through the violation of sum rules: for our crystal Hamiltonian in Eq. 2.35, the following commutator identity is trivially verified

$$[\hat{r}^\beta, [\hat{r}^\alpha, \hat{H}_0]] = -\frac{\hbar^2}{m}\,\delta^{\beta\alpha} \tag{2.76}$$

but such identities require a sum over an infinite number of bands. If some intermediate states are removed, the identity no longer holds. Since Eq. 2.76 is employed in the manipulations that bring the expressions from the length to the velocity gauge

form, one can begin to envision how their equivalence is lost when truncating the Hilbert space. This is the argument given originally in [48, 49].

When is band truncation appropriate? This is difficult to answer due to the integration over the FBZ, but we can attempt an estimation [57]. Transitions to higher energy bands can be removed if their contribution to the summation in the nonlinear conductivity is negligible. If we imagine starting with a given unperturbed Hamiltonian H_0 and deforming it by increasing the energy separation Δ between our subspace of interest and its complement (while maintaining the eigenstates unchanged), then intuitively we expect that at some point the contributions from outside our subspace will be too far off in energy to be relevant. This should indeed be true if the perturbation is much smaller than Δ. For the length gauge, a rough estimation is that this would be the case for $e\,\mathcal{A}^\alpha E^\alpha(\omega)/\Delta \ll 1$. The condition is more restrictive and harder to justify in the velocity gauge: $(e\,p^\alpha E^\alpha(\omega)/m\,\omega)/\Delta \simeq e\,\mathcal{A}^\alpha \Delta\, E^\alpha(\omega)/\omega\,\Delta = e\,\mathcal{A}^\alpha\, E^\alpha(\omega)/\omega \ll 1$, which always fails for sufficiently small frequencies ω. We conclude then, that if both of the previous conditions are satisfied (e. g. by having very small momentum matrix elements), band truncation is appropriate for both gauges and their results should remain approximately the same.

There is, however, no guarantee that this will be the case and in the velocity gauge the approximation breaks down in the DC limit, when all bands are required for accurate computation. Indeed, the velocity gauge is known to give wrong answers. Due to the ω^{-1} factors, the nonlinear conductivity appears to diverge in the limit $\omega \to 0$, even for the case of insulators, where no carriers are present and the response must physically be zero. It is zero in the length gauge and, by gauge invariance, so must be in the velocity gauge, with some cancellations making it that the conductivity only *appears* to diverge. But once a finite number of bands is considered and the equivalence between gauges is broken, the velocity gauge approach is the one that suffers the most, as the delicate cancellations that exist in the DC limit fail and the results are nonsense. Unphysical infrared divergences plague the velocity gauge, rendering it inoperable for most realistic calculations.

In truth, one rarely starts with a many band calculation that is later truncated to an appropriate subspace. More often, an effective Hamiltonian is defined in that subspace as the starting point and the velocity gauge expressions are then utterly inappropriate. The reasons for why the minimal coupling Hamiltonian is so much more sensitive to approximations are discussed further in Chapter 4.

At this point, it should be clear that despite being equivalent, each choice of gauge has its advantages and disadvantages, as it usual with gauge fixing. To close this discussion off, it is helpful to summarize them for both the length and velocity gauge perspectives, as well as consider when each would be preferable.

The length gauge is, currently, the standard for computations of the nonlinear conductivity. The perturbation breaks the translation symmetry of the lattice, causing transitions between neighbouring $\mathbf{k}$-points in the FBZ through a covariant derivative. The presence of covariant differentiation is responsible for more complicated expressions for the nonlinear conductivities, with a complexity that increases dramatically with the order of the perturbation. Still, the formalism is solid and the results well behaved, with no unphysical infrared divergences. Because of versatility of these expressions, that work well for finite band models, this approach is to be favored when pursuing analytical answers using effective Hamiltonians.

Contrarily, the velocity gauge preserves translation symmetry and no derivatives are involved, other than those required for the momentum matrix elements. The energy denominators involve simple poles only and the contributions from each point in the FBZ are independent. For this reason, it is the approach to use in numerical computations using tight-binding and other band structure models that provide bands defined over the entirety of the FBZ. As it stands, spurious divergences plague the velocity gauge formulation, but these problems are resolved in Chapter 4.

Before ending our discussion of the perturbation theory behind a typical calculation of the nonlinear optical conductivity, there is, alas, yet one more issue to cope with. A careful look at the denominators in previous equations raises the question: what happens when $\omega = \Delta\epsilon$? In other words, what happens under resonance conditions? This is particularly relevant in solids, where, once the optical frequency is above the gap, the resonance condition will be always met for some points in the FBZ. The truth is that most of the expressions presented so far are not strictly physical, nor do they make much mathematical sense, since they always diverge for resonant frequencies. This can be resolved by extending the frequencies into the Argand plane.

2.5 Complex frequencies

A careful inspection of the formulas derived in previous sections raises some subtle issues. For any frequency ω for which a resonance exists, the denominators in Eq. 2.29 are zero and the conductivity is strictly undefined. This difficulty traces back to Eq. 1.11 where the existence of a frequency domain nonlinear conductivity relies on the convergence of the Fourier transform. When ignoring any kind of relaxation mechanism, the response to an impulse given at an instant of time can last forever: $\sigma(t = +\infty) \neq 0$ and the Fourier transform diverges. This problem can be circumvented by extending the definition in Eq. 1.11 to complex frequencies in the upper half-plane [3],

$$\begin{aligned}
&\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1, \ldots, \bar{\omega}_n) \\
&= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sigma^{\beta\alpha_1\ldots\alpha_n}(t_1, \ldots, t_n)\, e^{i\bar{\omega}_1 t_1} \ldots e^{i\bar{\omega}_n t_n}\, dt_1 \ldots dt_n \\
&= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \left(\sigma^{\beta\alpha_1\ldots\alpha_n}(t_1, \ldots, t_n)\, e^{-\gamma t_1} \ldots e^{-\gamma t_n}\right) e^{i\omega_1 t_1} \ldots e^{i\omega_n t_n}\, dt_1 \ldots dt_n
\end{aligned} \tag{2.77}$$

with $\omega = \mathrm{Re}(\bar{\omega})$ and $\gamma = \mathrm{Im}(\bar{\omega})$.

The extension to complex frequencies can be interpreted in two ways. One is to consider the response function in Eq. 2.77 to be associated not to monochromatic waves, but to fields that are adiabatically switched on from the infinite past $E(\omega)\, e^{-i\omega t}\, e^{\gamma t}$.

A different perspective is to look at complex frequencies as a simple phenomenological method to introduce relaxation into the system. As stated in Eq. 2.77, the nonlinear conductivity in the frequency domain $\sigma'(\omega_1, \ldots, \omega_n) \equiv \sigma(\bar{\omega}_1, \ldots, \bar{\omega}_n)$ is obtained from a Fourier transform of a time domain response function that has the form $\sigma'(t_1, \ldots, t_n) = \sigma(t_1, \ldots, t_n)\, e^{-\gamma t_1} \ldots e^{-\gamma t_n}$ and satisfies the condition $\sigma'(t_i = \infty) = 0$.

This approach to relaxation is most certainly too simplistic to properly account for all the possible relaxation mechanisms that are observed in experiments, but it provides a direct and easy way to obtain sensible answers and it has advantages relative to the traditional approach of adding an phenomenological term to the equation of motion (Eq. 2.1) [56, 62]. For simplicity, we here take the imaginary part of the frequencies to be a constant γ, but more generally we could have $\gamma = \gamma(\omega)$. It would only be required that the function $\gamma(\omega)$ be even, in order for the reality condition[7] to be maintained.

When the relaxation-free limit is considered, where the imaginary parts of the frequencies are taken to zero from above, the integrand in Eq. 2.29 can always be defined as a distribution by making use of the Sokhotski-Plemelj theorem,

$$\int \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{g_{\mathbf{k}}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}}} \xrightarrow{\gamma\to 0^+} \fint \frac{d^d\mathbf{k}}{(2\pi)^d} \frac{g_{\mathbf{k}}}{\hbar\omega - \Delta\epsilon_{\mathbf{k}}} - \pi\, i \int \frac{d^d\mathbf{k}}{(2\pi)^d}\, g_{\mathbf{k}}\, \delta(\hbar\omega - \Delta\epsilon_{\mathbf{k}}) \quad (2.78)$$

For an atomic system, the distribution will be defined relative to an integral over the frequencies (Eq. 1.9) and relies on the condition that the spectral width of $E(\omega)$ be much greater than γ. For a crystal, the distribution is already accommodated by the presence of an integration over the FBZ and no restrictions must be applied to the optical fields considered. By taking the limit $\gamma \to 0^+$ in the expression[8]

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1, \ldots, \bar{\omega}_n) = e^n \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{J^{\beta}_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 + \cdots + \hbar\bar{\omega}_n - \Delta\epsilon_{\mathbf{k}ab}} \left[r^{\alpha_n}, \ldots \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}}} \circ [r^{\alpha_1}, \rho_0] \ldots \right]_{\mathbf{k}ab} \quad (2.79)$$

the nonlinear conductivity can be defined as a regular function for real frequencies and vanishing relaxation. It is implicitly assumed in this reasoning that no more than a single denominator diverges at a given $\mathbf{k}$. Otherwise, a singular nonlinear conductivity in the relaxation-free limit becomes a possibility[9].

In Eq. 2.79, the nonlinear conductivity can be further extended into the lower half-plane by analytic continuation [27]. In this way, Eq. 2.79 provides a valid expression over the entire complex frequency plane, even in regions where Eq. 2.77 no longer applies and the response function is, therefore, no longer physical.

[7] The time-domain conductivity must be real. See Eqs. 6.4 and 6.5.

[8] Same reasoning applies to the velocity gauge expressions.

[9] This does happen for certain frequency combinations and these are real physical divergences of the nonlinear conductivity (e. g. current injection).

Chapter 3

Two-band models in two dimensions

In the course of this thesis, there will be much discussion on the properties of the nonlinear conductivity and the techniques presented to compute it. To validate these, examples of physical systems with simple yet interesting nonlinear optical properties are required.

In this chapter, a popular condensed matter system is described that will serve this purpose: monolayer graphene. It is the first and most studied two-dimensional crystal ($d = 2$ in both real and reciprocal space). The lower dimensionality reduces the number of tensor components to be evaluated and alleviates the effort in the numerical integration over the FBZ. Hexagonal boron nitride crystals are also considered by adding a mass term to the graphene Hamiltonian, opening a gap and breaking the inversion symmetry of the graphene lattice. In introducing these systems, general modelling strategies are delineated, together with the process of retrieving the relevant information from them.

As far as the nonlinear optical properties are concerned, within the current formalism (Chapter 2), the specification of the band structure $\epsilon_{\mathbf{k}a}$ and the non-abelian Berry connection $\mathcal{A}_{\mathbf{k}ab}$ is all that matters and equates to defining the electronic system under study. In principle, these could be provided by any of the standard band structure theories[1]. In this text, we steer wide from sophisticated density functional theory calculations, which could perhaps provide more accurate results, but that are not necessary for our intent of studying response functions of systems of independent electrons. For us, intuitive, straightforward tight binding models will do the job.

3.1 The connection with tight binding

In a tight binding description [58], the Hilbert space is defined through a Wannier basis $\{\varphi_\lambda(\mathbf{r}-\mathbf{R}_n)\}$, with λ as the label that identifies the orbital inside the unit cell and $\mathbf{R}_n$ as a vector from the Bravais lattice. There is at least one state per atom in the crystal (an infinite number in the thermodynamic limit) and the wavefunction $\varphi_\lambda(\mathbf{r}-\mathbf{R}_n)$ is assumed to be well-localized around the position $\mathbf{R}_n + \boldsymbol{\delta}_\lambda$, where $\boldsymbol{\delta}_\lambda$

[1] It is perhaps worth mentioning that not only the band structure, but also some information on the stationary states must be provided by the model, so that the Berry connection may be calculated.

represents the position of the λ orbital relative to the unit cell center. This basis is orthogonal by construction,

$$\int d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r}-\mathbf{R}_m)\, \varphi_{\lambda'}(\mathbf{r}-\mathbf{R}_n) = \delta_{\lambda\lambda'}\, \delta_{\mathbf{R}_n,\mathbf{R}_m} \tag{3.1}$$

The tight binding Hamiltonian is derived by specifying the coupling between closely spaced Wannier orbitals in the crystal, as well as any relevant on-site energies. Often, nearest neighbour or next-to-nearest neighbour hoppings are sufficient to capture the physics.

$$\hat{H}_0 = \sum_{\lambda,\lambda'} \sum_{\mathbf{R}_n} \sum_{\mathbf{R}_m} |\varphi_\lambda(\mathbf{r}-\mathbf{R}_n)\rangle\; t_{\lambda\lambda'}(\mathbf{R}_n,\mathbf{R}_m)\; \langle\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}_m)| \tag{3.2}$$

where $t_{\lambda\lambda'}$ are the tight binding parameters and due to the translation symmetry of the lattice,

$$t_{\lambda\lambda'}(\mathbf{R}_n,\mathbf{R}_m) = t_{\lambda\lambda'}(\mathbf{R}_n-\mathbf{R}_m) \tag{3.3}$$

A second basis, which we shall name the sublattice Bloch basis, can be constructed by Fourier transform,

$$\Psi_{\mathbf{k}\lambda}(\mathbf{r}) = \sqrt{v_c} \sum_{\mathbf{R}_n} e^{i\mathbf{k}\cdot(\mathbf{R}_n+\boldsymbol{\delta}_\lambda)} \varphi_\lambda(\mathbf{r}-\mathbf{R}_n) \tag{3.4}$$

with $\mathbf{k}$ in the FBZ. It is also orthogonal,

$$\begin{aligned}
&\int d^3\mathbf{r}\, \Psi^*_{\mathbf{k}'\lambda'}(\mathbf{r})\, \Psi_{\mathbf{k}\lambda}(\mathbf{r}) \\
&\quad = v_c \sum_{\mathbf{R}_n,\mathbf{R}_m} e^{i\mathbf{k}\cdot(\mathbf{R}_n+\boldsymbol{\delta}_\lambda)} e^{-i\mathbf{k}'\cdot(\mathbf{R}_m+\boldsymbol{\delta}_{\lambda'})} \int d^3\mathbf{r}\, \varphi^*_{\lambda'}(\mathbf{r}-\mathbf{R}_m)\varphi_\lambda(\mathbf{r}-\mathbf{R}_n) \\
&\quad = v_c \sum_{\mathbf{R}_n,\mathbf{R}_m} e^{i\mathbf{k}\cdot(\mathbf{R}_n+\boldsymbol{\delta}_\lambda)}\, e^{-i\mathbf{k}'\cdot(\mathbf{R}_m+\boldsymbol{\delta}_{\lambda'})}\; \delta_{\lambda\lambda'}\, \delta_{\mathbf{R}_n,\mathbf{R}_m} \\
&\quad = v_c \sum_{\mathbf{R}_n} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_n}\; \delta_{\lambda\lambda'}\, e^{i(\mathbf{k}-\mathbf{k}')\cdot\boldsymbol{\delta}_\lambda} \\
&\quad = \delta_{\lambda\lambda'}\, (2\pi)^d\, \delta(\mathbf{k}-\mathbf{k}')
\end{aligned} \tag{3.5}$$

Rewriting the Hamiltonian in the sublattice Bloch basis will show $\mathbf{k}$ to be a good quantum number,

$$\hat{H}_0 = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{\lambda,\lambda'} |\Psi_{\mathbf{k}\lambda}\rangle\, (H_0)_{\mathbf{k}\lambda\lambda'}\, \langle\Psi_{\mathbf{k}\lambda'}| \tag{3.6}$$

with

$$(H_0)_{\mathbf{k}\lambda\lambda'} = \sum_{\mathbf{R}_n} t_{\lambda\lambda'}(\mathbf{R}_n)\, e^{-i\mathbf{k}\cdot\mathbf{R}_n}\, e^{-i\mathbf{k}\cdot(\boldsymbol{\delta}_\lambda-\boldsymbol{\delta}_{\lambda'})} \tag{3.7}$$

The basis functions in Eq. 3.4 are indeed Bloch functions, with the form of Eq. 2.36, but are not, usually, the eigenstates of the tight binding Hamiltonian. The

Hamiltonian can, however, be diagonalized and the stationary states represented in the sublattice Bloch basis,

$$\psi_{\mathbf{k}a}(\mathbf{r}) = \sum_{\lambda} c_{\mathbf{k}a\lambda}\, \Psi_{\mathbf{k}\lambda}(\mathbf{r}) \tag{3.8}$$

The energy eigenstates must also form an orthogonal basis,

$$\begin{aligned}\int d^3\mathbf{r}\; \psi^*_{\mathbf{k}'a}(\mathbf{r})\psi_{\mathbf{k}b}(\mathbf{r}) &= \sum_{\lambda\lambda'} c^*_{\mathbf{k}'a\lambda'} c_{\mathbf{k}b\lambda} \int d^3\mathbf{r}\; \Psi^*_{\mathbf{k}'\lambda'}(\mathbf{r})\Psi_{\mathbf{k}\lambda}(\mathbf{r}) \\ &= (2\pi)^d\, \delta(\mathbf{k}-\mathbf{k}') \sum_{\lambda} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda} \\ &= (2\pi)^d\, \delta(\mathbf{k}-\mathbf{k}')\, \delta_{ab}\end{aligned} \tag{3.9}$$

Orthogonality of $\psi_{\mathbf{k}a}(\mathbf{r})$ translates into the following condition on the coefficients of Eq. 3.8,

$$\sum_{\lambda} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda} = \delta_{ab} \tag{3.10}$$

which just means that the representation of the energy eigenstates $\psi_{\mathbf{k}a}(\mathbf{r})$ in the sublattice Bloch basis is given by orthogonal vectors.

Having all the proper definitions and normalization issues settled, we can proceed to our quantities of interest. The eigenvalues associated to $\psi_{\mathbf{k}a}$ give the band structure,

$$\sum_{\lambda'} (H_0)_{\mathbf{k}\lambda\lambda'}\, c_{\mathbf{k}a\lambda'} = \epsilon_{\mathbf{k}a}\, c_{\mathbf{k}a\lambda} \tag{3.11}$$

and from the energy eigenstates, we calculate the non-abelian Berry connection,

$$\begin{aligned}\mathcal{A}^{\alpha}_{\mathbf{k}ab} &= i\,\langle u_{\mathbf{k}a}|\partial^{\alpha} u_{\mathbf{k}b}\rangle = \frac{i}{v_c}\int_{uc} d^3\mathbf{r}\; u^*_{\mathbf{k}a}(\mathbf{r})\,(\partial^{\alpha} u_{\mathbf{k}b}(\mathbf{r})) \\ &= \frac{i}{v_c}\int_{uc} d^3\mathbf{r}\; e^{i\,\mathbf{k}\cdot\mathbf{r}}\,\psi^*_{\mathbf{k}a}(\mathbf{r})\left(\partial^{\alpha}\left(e^{-i\,\mathbf{k}\cdot\mathbf{r}}\,\psi_{\mathbf{k}b}(\mathbf{r})\right)\right) \\ &= \frac{i}{v_c}\sum_{\lambda,\lambda'}\int_{uc} d^3\mathbf{r}\; e^{i\mathbf{k}\cdot\mathbf{r}} c^*_{\mathbf{k}a\lambda}\,\Psi^*_{\mathbf{k}\lambda}(\mathbf{r})\left(\partial^{\alpha}\left(e^{-i\mathbf{k}\cdot\mathbf{r}} c_{\mathbf{k}b\lambda'}\,\Psi_{\mathbf{k}\lambda'}(\mathbf{r})\right)\right) \\ &= i\sum_{\lambda,\lambda'}\sum_{\mathbf{R}_n,\mathbf{R}_m}\int_{uc} d^3\mathbf{r}\; e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R}_n-\boldsymbol{\delta}_{\lambda})} c^*_{\mathbf{k}a\lambda}\,\varphi^*_{\lambda}(\mathbf{r}-\mathbf{R}_n) \\ &\qquad\qquad \times\left(\partial^{\alpha}\left(e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R}_m-\boldsymbol{\delta}_{\lambda'})} c_{\mathbf{k}b\lambda'}\,\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}_m)\right)\right)\end{aligned} \tag{3.12}$$

The derivative will act on two factors: the coefficient $c_{\mathbf{k}b\lambda'}$ and the phase factor. The first term,

$$
\begin{aligned}
& i\sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda'}) \sum_{\mathbf{R}_n,\mathbf{R}_m} \int_{uc} d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r}-\mathbf{R}_n)\varphi_\lambda(\mathbf{r}-\mathbf{R}_m) e^{i\mathbf{k}\cdot(\mathbf{R}_m-\mathbf{R}_n)} e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \\
=& i\sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda'}) \sum_{\mathbf{R}_n,\mathbf{R}'_m} \int_{uc} d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r}-\mathbf{R}_n)\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}'_m-\mathbf{R}_n) e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \\
=& i\sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda'}) \sum_{\mathbf{R}'_m} e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \int d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r})\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}'_m) \\
=& i\sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda'}) \sum_{\mathbf{R}'_m} e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, \delta_{\lambda\lambda'}\, \delta_{\mathbf{R}'_m,\mathbf{0}} \\
=& i\sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda'})\, \delta_{\lambda\lambda'} \\
=& i\sum_{\lambda} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda})
\end{aligned}
\tag{3.13}
$$

The second term,

$$
\begin{aligned}
& \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda'} \sum_{\mathbf{R}_n,\mathbf{R}_m} \int_{uc} d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r}-\mathbf{R}_n)\,(r^\alpha - R^\alpha_n - \delta^\alpha_{\lambda'})\,\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}_m)\, e^{i\mathbf{k}\cdot(\mathbf{R}_m-\mathbf{R}_n)} e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \\
=& \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda'} \sum_{\mathbf{R}_n,\mathbf{R}'_m} \int_{uc} d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r}-\mathbf{R}_n)\left(r^\alpha - R^\alpha_n - R'^\alpha_m - \delta^\alpha_{\lambda'}\right)\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}'_m-\mathbf{R}_n)\, e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \\
=& \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda'} \sum_{\mathbf{R}'_m} e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \int d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r})\left(r^\alpha - R'^\alpha_m - \delta^\alpha_{\lambda'}\right)\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}'_m) \\
=& \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda'} \sum_{\mathbf{R}'_m} e^{i\mathbf{k}\cdot\mathbf{R}'_m}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)}\, S^\alpha_{\lambda\lambda'}(\mathbf{R}'_m) \\
=& \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda} c_{\mathbf{k}b\lambda'}\, S^\alpha_{\mathbf{k}\lambda\lambda'}
\end{aligned}
\tag{3.14}
$$

where the last two steps were mere definitions:

$$
\mathbf{S}_{\lambda\lambda'}(\mathbf{R}_n) \equiv \int d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r})\,(\mathbf{r}-\mathbf{R}_n-\boldsymbol{\delta}_{\lambda'})\,\varphi_{\lambda'}(\mathbf{r}-\mathbf{R}_n) \tag{3.15}
$$

$$
\mathbf{S}_{\mathbf{k}\lambda\lambda'} \equiv \sum_{\mathbf{R}_n} \mathbf{S}_{\lambda\lambda'}(\mathbf{R}_n)\, e^{i\mathbf{k}\cdot\mathbf{R}_n}\, e^{i\mathbf{k}\cdot(\boldsymbol{\delta}_{\lambda'}-\boldsymbol{\delta}_\lambda)} \tag{3.16}
$$

Assembling the two terms, we have the general form of the Berry connection,

$$
\mathcal{A}^\alpha_{\mathbf{k}ab} = i\sum_{\lambda} c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda}) + \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}\, c_{\mathbf{k}b\lambda'}\, S^\alpha_{\mathbf{k}\lambda\lambda'} \tag{3.17}
$$

This is a more general expression than the one typically employed in the literature. We can take the limit were the Wannier states are not only orthogonal, but so localized that no overlap exists between them. In that case[2],

[2]Note that this approximation becomes harder to justify for models with more than one orbital per atom. For such systems, there is surely overlap between orbitals and the approximation equates to assuming that, for whatever reason, the following integral is negligible: $\int d^3\mathbf{r}\, \varphi^*_\lambda(\mathbf{r})\,\mathbf{r}\,\varphi_{\lambda'}(\mathbf{r}) \approx 0$, with $\lambda \neq \lambda'$.

$$\mathbf{S}_{\lambda\lambda'}(\mathbf{R}_n) = \delta_{\lambda\lambda'}\,\delta_{\mathbf{R}_n,0}\,\mathbf{S}_{\lambda\lambda'}(0) \tag{3.18}$$

where

$$\mathbf{S}_{\lambda\lambda}(0) = \left(\int d^3\mathbf{r}\;|\varphi_\lambda(\mathbf{r})|^2\,\mathbf{r}\right) - \boldsymbol{\delta}_\lambda = 0 \tag{3.19}$$

since $\boldsymbol{\delta}_\lambda$ stands for the average position of the orbital λ in the unit cell. In the limit of no overlap, the Berry connection simplifies to

$$\mathcal{A}^\alpha_{\mathbf{k}ab} = i\sum_\lambda c^*_{\mathbf{k}a\lambda}\,(\partial^\alpha c_{\mathbf{k}b\lambda}) \tag{3.20}$$

This is the formula we shall use to compute $\mathcal{A}$ from tight binding models.

It is worth pointing out that the notion of a Berry *connection* implies the existence of an Hamiltonian defined over a space of parameters, over which $\mathcal{A}$ could be integrated to give the Berry phase [59]. In band structure theories, the parameter is the Bloch vector $\mathbf{k}$ and the parameter space is the First Brillouin Zone. The parametric Hamiltonian $\hat{H}_0(\mathbf{k})$ is defined[3] through the matrix elements in Eq. 3.7. Its eigenvalues provide the band structure and the eigenstates, through Eq. 3.20, permit us to compute the non-abelian Berry connection.

$\hat{H}_0(\mathbf{k})$ can have any dimension, depending on the number of orbitals per unit cell. For simplicity, we will focus on systems with a single orbital per atom and two atoms per unit cell. That is to say, on two-band models.

3.2 Berryology of the two-band model

Any two-band model can be written as a combination of Pauli matrices,

$$\hat{H}_0(\mathbf{k}) = d_0(\mathbf{k}) + \hat{\boldsymbol{\sigma}}\cdot\boldsymbol{d}(\mathbf{k}) \tag{3.21}$$

with $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, for some functions $\boldsymbol{d}(\mathbf{k}) = (d_x(\mathbf{k}), d_y(\mathbf{k}), d_z(\mathbf{k}))$.

This Hamiltonian could be derived from tight binding as described before or from a low energy effective theory, valid only in a restricted region of the FBZ, or by any other means by which a two-band description of the electron dynamics could be constructed. In this subsection, we derive the band structure and the Berry connection for the generic two-band model in Eq 3.21.

The study of the two-band model is of special importance. The two-band model is the condensed matter analogue of the two-level atom, providing as simple a model as possible, making analytical calculations feasible and allowing the concepts to emerge clearly, while still being complex enough to merit attention and to have a wide range of applicability, describing any situation where incident photon frequencies connect a single pair of bands around the Fermi surface. A still simpler model is the single band model, which is appropriate to describe metallic systems at low frequencies, but permits only intraband transitions and is a bit too simple for our goals here. The two-band model contains combinations of intra- and interband transitions that make the dynamics of the electron under optical excitation nontrivial and introduces most of the key concepts relevant for more complicated multiband systems.

[3] See Section 4.5 for further details.

A particular Hamiltonian can be specified through the d_0 and $\boldsymbol{d}$ functions. The first term $d_0(\mathbf{k})$ in Eq. 3.21 merely shifts the energies (in a $\mathbf{k}$-dependent fashion). Likewise, the modulus of $\boldsymbol{d}(\mathbf{k})$ determines the band structure, but has no impact on the nature of the eigenstates[4]. The stationary states, and therefore the Berry connection, are a function only of the *direction* of $\boldsymbol{d}(\mathbf{k})$. All geometrical properties of the system fall on this versor field $\hat{\boldsymbol{d}}(\mathbf{k})$, that can be seen as a map between the FBZ and the 2-sphere S^2,

$$d_0(\mathbf{k})\,\mathbb{1} + \begin{pmatrix} d_z(\mathbf{k}) & d_x(\mathbf{k}) - i\,d_y(\mathbf{k}) \\ d_x(\mathbf{k}) + i\,d_y(\mathbf{k}) & -d_z(\mathbf{k}) \end{pmatrix} \tag{3.22}$$

where

$$\boldsymbol{d} = (d_x, d_y, d_z) = d\,(\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi) \tag{3.23}$$

with the $\mathbf{k}$ dependence henceforth left implicit.

The band structure is given by

$$\epsilon_{\mathbf{k}c} = d_0 + |\boldsymbol{d}| \qquad \epsilon_{\mathbf{k}v} = d_0 - |\boldsymbol{d}| \tag{3.24}$$

with the subscripts c and v standing for conduction and valence bands, respectively, while the eigenstates take the form

$$c_{\mathbf{k}c\lambda} \to \begin{pmatrix} \cos(\phi/2) \\ \sin(\phi/2)e^{i\theta} \end{pmatrix} \qquad c_{\mathbf{k}v\lambda} \to \begin{pmatrix} \sin(\phi/2) \\ -\cos(\phi/2)e^{i\theta} \end{pmatrix} \tag{3.25}$$

Notice that, for this particular choice of gauge, the states in the conduction band are uniquely defined everywhere[5] on the 2-sphere with the exception of the south pole ($\phi = \pi$), while in the valence band the states are only multivalued at the north pole ($\phi = 0$). The same will be true for the abelian Berry connection of the respective bands (the diagonal matrix elements of $\mathcal{A}_{\mathbf{k}ab}$).

The non-abelian Berry connection is computed from the states in Eqs. 3.25 with Eq. 3.20,

$$\boldsymbol{\mathcal{A}}_{\mathbf{k}ab} = \begin{pmatrix} -\nabla\theta\,\sin^2(\phi/2) & \frac{1}{2}\nabla\theta\sin\phi + \frac{i}{2}\nabla\phi \\ \frac{1}{2}\nabla\theta\sin\phi - \frac{i}{2}\nabla\phi & -\nabla\theta\,\cos^2(\phi/2) \end{pmatrix}_{ab} \tag{3.26}$$

The previous results can be rephrased without reference to the spherical coordinates, by writing everything explicitly in terms of the $\boldsymbol{d}$ field. Inverting Eq. 3.23, it is straightforward to write the angles and their derivatives as a function of $\boldsymbol{d}$,

$$\theta = \arctan\left(\frac{d_y}{d_x}\right) \qquad \phi = \arccos\left(\frac{d_z}{d}\right) \tag{3.27}$$

$$\nabla\theta = \frac{d_x\,\nabla d_y - d_y\,\nabla d_x}{d_x^2 + d_y^2} \qquad \nabla\phi = \frac{d_z\,\nabla d - d\,\nabla d_z}{d\sqrt{d_x^2 + d_y^2}} \tag{3.28}$$

$$\sin\phi = \frac{\sqrt{d_x^2 + d_y^2}}{d} \qquad \cos(\phi/2) = \sqrt{\frac{d + d_z}{2d}} \qquad \sin(\phi/2) = \sqrt{\frac{d - d_z}{2d}} \tag{3.29}$$

[4] Aside from the possibility of degeneracy when $|\boldsymbol{d}| = 0$.

[5] The statement "the states are uniquely defined" refers to, of course, the kets representing the physical states and the phase factors they carry.

Replacing back in Eqs. 3.25 and 3.26, we rewrite the eigenstates

$$c_{\mathbf{k}c\lambda} \to \frac{1}{\sqrt{2d}} \begin{pmatrix} \sqrt{d+d_z} \\ \frac{d_x + i\, d_y}{\sqrt{d+d_z}} \end{pmatrix} \qquad c_{\mathbf{k}v\lambda} \to \frac{1}{\sqrt{2d}} \begin{pmatrix} \sqrt{d-d_z} \\ -\frac{d_x + i\, d_y}{\sqrt{d-d_z}} \end{pmatrix} \tag{3.30}$$

and the Berry connection

$$\boldsymbol{\mathcal{A}}_{\mathbf{k}ab} = \begin{pmatrix} -\frac{d_x\, \nabla d_y - d_y\, \nabla d_x}{d_x^2 + d_y^2} \left(\frac{d - d_z}{2d}\right) & \frac{(d_x\, \nabla d_y - d_y\, \nabla d_x) + i\, (d_z\, \nabla d - d\, \nabla d_z)}{2\, d\, \sqrt{d_x^2 + d_y^2}} \\ \frac{(d_x\, \nabla d_y - d_y\, \nabla d_x) - i\, (d_z\, \nabla d - d\, \nabla d_z)}{2\, d\, \sqrt{d_x^2 + d_y^2}} & -\frac{d_x\, \nabla d_y - d_y\, \nabla d_x}{d_x^2 + d_y^2} \left(\frac{d + d_z}{2d}\right) \end{pmatrix}_{ab} \tag{3.31}$$

entirely in terms of the functions d_x, d_y and d_z.

As mentioned before, it is not possible to define the states uniquely over the 2-sphere. Hence, there may be regions of the FBZ where the Berry connection is left undefined for a given choice of gauge. This is because, as addressed in Section 2.3, the Berry connection is not invariant under a $U(1)$ gauge transformation on the Bloch states.

Despite this, physical quantities like the nonlinear conductivity cannot depend on the particular choice of phase in the Bloch functions and must be built of gauge invariant objects. Notably, the Berry curvature is known to be invariant under such gauge transformations. It is defined by

$$\mathcal{F}_a^{\alpha\beta} \equiv \partial^\alpha \mathcal{A}^\beta_{\mathbf{k}aa} - \partial^\beta \mathcal{A}^\alpha_{\mathbf{k}aa} \tag{3.32}$$

It can be obtained from the abelian Berry connection $\mathcal{A}_{\mathbf{k}aa}$ in Eq. 3.26. An often simpler approach, however, follows from rewriting the curvature in terms of the off-diagonal matrix elements of $\boldsymbol{\mathcal{A}}$,

$$\mathcal{F}_a^{\alpha\beta} = \partial^\alpha \mathcal{A}^\beta_{\mathbf{k}aa} - \partial^\beta \mathcal{A}^\alpha_{\mathbf{k}aa} = i \left[\mathcal{A}^\alpha, \mathcal{A}^\beta\right]_{\mathbf{k}aa} = i \sum_b \left(\mathcal{A}^\alpha_{\mathbf{k}ab} \mathcal{A}^\beta_{\mathbf{k}ba} - \mathcal{A}^\beta_{\mathbf{k}ab} \mathcal{A}^\alpha_{\mathbf{k}ba}\right) \tag{3.33}$$

where the commutation property of position operators (covariant derivatives), discussed in Appendix C, was used. Eq. 3.33 is closely related to the well-known rewriting of the Berry curvature in terms of derivatives of the Hamiltonian and energy differences [59]. The off-diagonal matrix elements also have the advantage that under a gauge transformation they only pick up a phase,

$$\mathcal{A}_{\mathbf{k}ab} \to \mathcal{A}_{\mathbf{k}ab}\, e^{i(\phi_{\mathbf{k}a} - \phi_{\mathbf{k}b})} \tag{3.34}$$

In three dimensions, the curvature tensor can be mapped into a vector field $\boldsymbol{\Omega}$,

$$\Omega_a^\alpha \equiv \frac{1}{2} \epsilon^{\alpha\beta\gamma} \mathcal{F}_a^{\beta\gamma} = \frac{i}{2} \sum_b \epsilon^{\alpha\beta\gamma} \left(\mathcal{A}^\beta_{\mathbf{k}ab} \mathcal{A}^\gamma_{\mathbf{k}ba} - \mathcal{A}^\gamma_{\mathbf{k}ab} \mathcal{A}^\beta_{\mathbf{k}ba}\right) = i \sum_b \left(\boldsymbol{\mathcal{A}}_{\mathbf{k}ab} \times \boldsymbol{\mathcal{A}}_{\mathbf{k}ba}\right)^\alpha \tag{3.35}$$

which is more easily visualized.

From Eqs. 3.26 and 3.33,

$$\begin{aligned}\mathcal{F}_c^{\alpha\beta} &= i\left(\left(\frac{\partial^\alpha\theta}{2}\sin\phi + i\,\frac{\partial^\alpha\phi}{2}\right)\left(\frac{\partial^\beta\theta}{2}\sin\phi - i\,\frac{\partial^\beta\phi}{2}\right) - (\alpha\leftrightarrow\beta)\right)\\ &= \frac{1}{2}\left(\partial^\alpha\theta\;\partial^\beta\phi - \partial^\beta\theta\;\partial^\alpha\phi\right)\sin\phi\end{aligned} \tag{3.36}$$

Similarly for the valence band,

$$\mathcal{F}_v^{\alpha\beta} = -\frac{1}{2}\left(\partial^\alpha\theta\;\partial^\beta\phi - \partial^\beta\theta\;\partial^\alpha\phi\right)\sin\phi \tag{3.37}$$

For the curvature vector field, it follows from Eqs. 3.26 and 3.35

$$\boldsymbol{\Omega}_c = \frac{1}{2}\left(\nabla\theta\times\nabla\phi\right)\sin\phi \tag{3.38}$$

This is the general form of the Berry curvature field for a two band model, written in spherical coordinates. Just as it was the case for the Berry connection, the curvature tensor can be expressed directly in terms of $\hat{\boldsymbol{d}}(\mathbf{k})$,

$$\mathcal{F}_c^{\alpha\beta} = -\mathcal{F}_v^{\alpha\beta} = -\frac{1}{2\,d^3}\,\boldsymbol{d}\cdot\left(\partial^\alpha\boldsymbol{d}\times\partial^\beta\boldsymbol{d}\right) \tag{3.39}$$

Finally, replacing Eqs. 3.28 and 3.29 in Eq. 3.38,

$$\boldsymbol{\Omega}_c = -\frac{1}{2d^3}\left(\left(\nabla d_x\times\nabla d_y\right)\,d_z + \left(\nabla d_y\times\nabla d_z\right)\,d_x + \left(\nabla d_z\times\nabla d_x\right)\,d_y\right) \tag{3.40}$$

Comparing with Eq. 3.38, we note that the right hand side of Eqs. 3.39 and 3.40 must still depend only on the direction of the $\boldsymbol{d}$ field[6], even if in practice it might be easier to use to non-normalized vectors in the calculations.

In studies of topology, the curvature is the central object of interest. In nonlinear optical response, many other gauge invariant objects exist that must be considered. A particularly simple example is the product $\mathcal{A}_{\mathbf{k}cv}\,\mathcal{A}_{\mathbf{k}vc}$, with more complicated quantities emerging with increasing order in perturbation theory. This will become clear in Chapter 5, when the nonlinear optical response of the general two-band model is detailed up to third order.

For the remainder of this chapter, the formulas in Eqs. 3.24, 3.25 and 3.26 are applied to specific examples.

3.3 Monolayer graphene

Two dimensional crystals are today on the frontier of condensed matter physics research [63]. The breakthrough that launched this field came in 2004, when the Manchester group isolated and studied mono- and few-layer graphene [64]. The astonishing electronic quality demonstrated by these materials immediately caught the attention of the community.

[6] It can also be proved by noticing that the right hand side of Eq. 3.39 has the form of a determinant.

Graphene is a two dimensional allotrope of carbon. It consist of carbon atoms displayed in a honeycomb pattern and, in a sense, forms the basis for all other allotropes of carbon [65]. Electrons moving in the honeycomb lattice behave as massless Dirac fermions, moving as if they were in the ultrarelativistic limit, but with a two orders of magnitude lower effective speed of light [65]. This unusual dispersion law of electrons in graphene is at the heart of most graphene physics and of the variety of amazing new properties it presents. Among these are the ballistic electronic transport [64], a new form of quantum hall effect and the possibility to observe it at room temperature [66], record high thermal conductivity [67], ...

Phenomena from high-energy physics are seen in table-top experiments in graphene and quantum phenomena typical of low temperature behaviour are observed in graphene at room temperature [65]. Beyond basic physics, graphene is an extreme case of surface science (its all surface!) and its chemistry is of considerable interest [68]. The carrier concentration in graphene can be controlled by applying an external voltage, adding tunability to the unique advantages of this system. Applications of the various properties of graphene are a subject of intensive research and of considerable interest from industry, leading to rapid progress in production methods of graphene sheets [69].

The optical properties of graphene have also attracted great interest. In the absence of doping and near the Dirac point, the linear conductivity is frequency independent, which causes some interesting features such as a transparency in the visible defined by universal constants [70] and a finite conductivity in the limit of no charge carriers [71]. The literature on the nonlinear optical response of graphene is extensive and only some of the more relevant works are mentioned in this thesis. Mikhailov [72] proposed theoretically that the linear dispersion in graphene should lead to considerable optical nonlinearities. The first measurement of the nonlinear response was made by Hendry *et al* [73]. Aside from the early Boltzmann equation calculations [72, 74, 75], analytical and numerical computations of the third order conductivity of graphene have been performed [76, 77], using a choice of gauge that stands somewhere between the described length and velocity gauges[7]. Disagreements are found with experimental results, but these differ several order of magnitude between themselves[8] [76]. This is perhaps partly due to the strong dispersion of the response and a lack of knowledge of the Fermi level (carrier concentration) in most of these experimental studies. A notable exception is the work of Jiang *et al* [78], where the sample was properly characterized and the nonlinear optical response measured for a wide range of doping levels, finding good agreement with theoretical predictions [76–78].

In this section, the electronic properties of graphene will be discussed at a basic level, with the intent of obtaining the band structure $\epsilon_{\mathbf{k}a}$ and the Berry connection $\mathcal{A}_{\mathbf{k}ab}$ required for calculations of the nonlinear conductivity.

[7]This hybrid gauge choice is obtained by performing an unitary transformation $\mathcal{U}(t) = \exp(-i\,e\,\hat{r}^{\alpha}_{diag}\,A^{\alpha}(t)/\hbar) = \exp(e\,(\hat{\partial}^{\alpha} - i\,\hat{\mathcal{A}}^{\alpha}_{diag})\,A^{\alpha}(t)/\hbar)$ on the length gauge Hamiltonian (Eq. 1.13) that involves the abelian Berry connection $\mathcal{A}_{diag}$ (it excludes the off-diagonal matrix elements of the position operator).

[8]See supplementary material of [78].

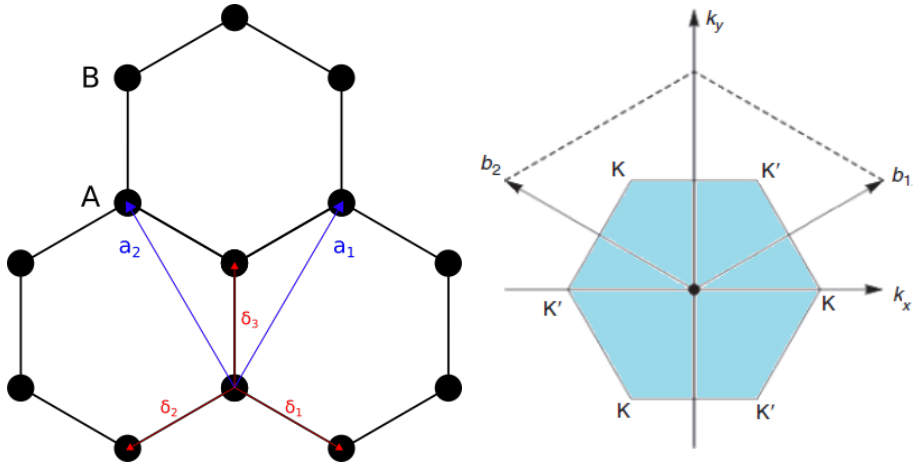


Figure 3.1: In the left, the honeycomb lattice of graphene. Identified in the figure are the primitive vectors of the Bravais lattice and the two sites, A and B, linked by the vectors $\boldsymbol{\delta}_i$; in the right, the First Brillouin Zone (FBZ) of graphene with the primitive vectors of the reciprocal lattice. The vertices of the FBZ are the Dirac points and all vertices can be obtained from only two by translations of the reciprocal lattice.

3.3.1 Honeycomb lattice

Although the carbon atoms are arranged in a honeycomb pattern in graphene, the Bravais lattice, which describes the translation symmetry of the crystal, is a triangular lattice. This can be seen by noting that there are two different kinds of sites in a honeycomb lattice, A and B sites, as depicted in Fig. 3.1. A translation from A to B is not a symmetry, since it does not leave the lattice invariant. The Bravais lattice has then two atoms per unit cell and is generated by the set of translations $n\,\mathbf{a}_1 + m\,\mathbf{a}_2$ for any integers n and m and the primitive vectors

$$\mathbf{a}_1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) a \qquad \mathbf{a}_2 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right) a \tag{3.41}$$

where $a = \sqrt{3}a_0$ is the lattice parameter and a_0 is the nearest neighbour distance (in graphene: $a_0 = 1.42$ Å).

The reciprocal lattice can be shown to be also a triangular lattice, with primitive vectors

$$\mathbf{b}_1 = \frac{4\pi}{\sqrt{3}a}\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right) \qquad \mathbf{b}_2 = \frac{4\pi}{\sqrt{3}a}\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right) \tag{3.42}$$

The First Brillouin Zone (FBZ) will be an hexagon (Fig. 3.1), with vertices at

$$\mathbf{K} = -\mathbf{K}' = \frac{\mathbf{b}_1 - \mathbf{b}_2}{3} = \frac{4\pi}{3a}(1, 0) \tag{3.43}$$

The other vertices of the hexagon can be obtained from $\mathbf{K}$ and $\mathbf{K}'$ with a translation by a reciprocal lattice element, and therefore correspond to the same Bloch states. These two states are referred to as Dirac points, for reasons to be made clear (Section 3.3.3).

The set of vectors linking an A-site atom to its neighbours, depicted in Fig. 3.1, is

$$\boldsymbol{\delta}_1 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) a_0 \qquad \boldsymbol{\delta}_2 = \left(-\frac{\sqrt{3}}{2}, -\frac{1}{2}\right) a_0 \qquad \boldsymbol{\delta}_3 = (0,1)\, a_0 \tag{3.44}$$

3.3.2 Tight binding model

A simple nearest neighbour tight binding model describes the basic properties of electrons in graphene. To construct the one-particle Hilbert space, one state (per spin) is assigned to each carbon atom. The basis states can then be written as

$$|\varphi_A(\mathbf{r}-\mathbf{R}_n)\rangle = |\varphi(\mathbf{r}-\mathbf{R}_n)\rangle \qquad |\varphi_B(\mathbf{r}-\mathbf{R}_n)\rangle = |\varphi(\mathbf{r}-\mathbf{R}_n-\boldsymbol{\delta}_3)\rangle \tag{3.45}$$

where we centered the origin of the unit cell in the A site, so that $\boldsymbol{\delta}_A = 0$ and $\boldsymbol{\delta}_B = \boldsymbol{\delta}_3$. These are the $\boldsymbol{\delta}_\lambda$ vectors introduced in Section 3.1, with the sublattice/orbital index taking on two possible values: $\lambda = A, B$.

There is, of course, more than one occupied orbital in a carbon atom. A more involved treatment shows that the four valence orbitals (per spin) combine to give three hybridized sp^2 orbitals in the graphene plane and one out-of-the-plane π orbital. The sp^2 orbitals form covalent bonds that are responsible for the strength and stability of graphene [65, 79]. The π states are responsible for the electronic transport in graphene [65, 79]. From the point of view of band theory, the sp^2 orbitals form deep valence bands, which are completely filled and don't contribute to electronic transport. We can then just consider the π states, which can be transformed into orthogonal Wannier states, as the ones in Eq. 3.45.

To keep it simple, only nearest neighbour interactions will be considered. Each Wannier state couples to the three neighbouring ones, giving rise to the following Hamiltonian,

$$\hat{H}_0 = t \sum_{\mathbf{R}_n} |\varphi_A(\mathbf{r}-\mathbf{R}_n)\rangle \left(\langle\varphi_B(\mathbf{r}-\mathbf{R}_n)| + \langle\varphi_B(\mathbf{r}-\mathbf{R}_n-\mathbf{a}_1)| + \langle\varphi_B(\mathbf{r}-\mathbf{R}_n-\mathbf{a}_2)|\right) + h.c. \tag{3.46}$$

where t is the tight binding parameter.

A Fourier transform will diagonalize this Hamiltonian. Following Section 3.1, the sublattice Bloch functions are introduced

$$\Psi_{\mathbf{k}A}(\mathbf{r}) = \sqrt{v_c} \sum_{\mathbf{R}_n} e^{i\mathbf{k}\cdot\mathbf{R}_n} \varphi_A(\mathbf{r}-\mathbf{R}_n) \qquad \Psi_{\mathbf{k}B}(\mathbf{r}) = \sqrt{v_c} \sum_{\mathbf{R}_n} e^{i\mathbf{k}\cdot(\mathbf{R}_n+\boldsymbol{\delta}_3)} \varphi_B(\mathbf{r}-\mathbf{R}_n) \tag{3.47}$$

and the Hamiltonian is rewritten in this basis. According to Eqs. 3.6 and 3.7,

$$\hat{H}_0 = t \int \frac{d^d\mathbf{k}}{(2\pi)^d}\, \Phi(\mathbf{k})\, |\Psi_{\mathbf{k}A}\rangle \langle\Psi_{\mathbf{k}B}| + h.c. \tag{3.48}$$

with

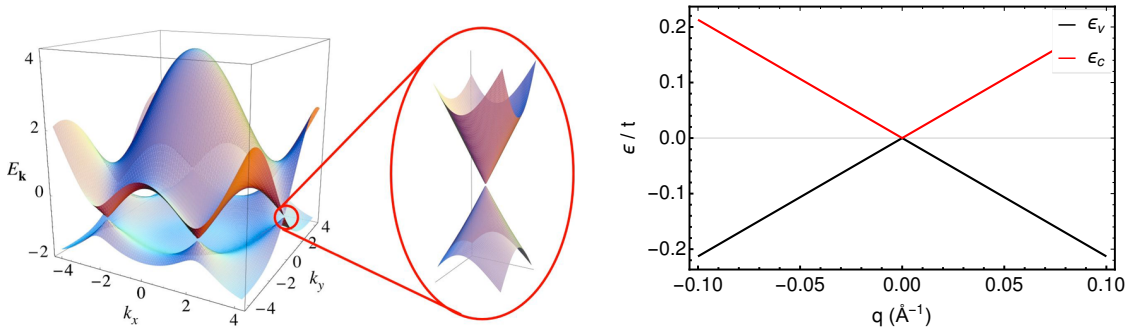


Figure 3.2: Band structure of monolayer graphene (π bands only). Zoomed in at the vertices of the FBZ, are the Dirac cones. In the left, the entire band structure over the FBZ (adapted from [65]). In the right, the linear dispersion near the Dirac point.

$$\Phi(\mathbf{k}) = |\Phi(\mathbf{k})|e^{-i\theta} = e^{i\mathbf{k}\cdot\boldsymbol{\delta}_3}\left(1 + e^{-i\mathbf{k}\cdot\mathbf{a}_1} + e^{-i\mathbf{k}\cdot\mathbf{a}_2}\right) = e^{i\mathbf{k}\cdot\boldsymbol{\delta}_1} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_2} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_3} \tag{3.49}$$

We are then left with a two-dimensional problem,

$$H_0 = t\begin{pmatrix} 0 & \Phi(\mathbf{k}) \\ \Phi^*(\mathbf{k}) & 0 \end{pmatrix} \tag{3.50}$$

This is our two-band model. Notice that the two states per $\mathbf{k}$ resulted from the existence of two atoms per unit cell. Comparing with Eq. 3.22, we have for graphene $d_0 = d_z = 0$ and

$$d_x = t\ \mathrm{Re}\{\Phi(\mathbf{k})\} = +t\cos(\mathbf{k}\cdot\boldsymbol{\delta}_1) + t\cos(\mathbf{k}\cdot\boldsymbol{\delta}_2) + t\cos(\mathbf{k}\cdot\boldsymbol{\delta}_3) \tag{3.51}$$

$$d_y = -t\ \mathrm{Im}\{\Phi(\mathbf{k})\} = -t\sin(\mathbf{k}\cdot\boldsymbol{\delta}_1) - t\sin(\mathbf{k}\cdot\boldsymbol{\delta}_2) - t\sin(\mathbf{k}\cdot\boldsymbol{\delta}_3) \tag{3.52}$$

In spherical coordinates (Eq. 3.23),

$$\phi = \frac{\pi}{2} \qquad \theta = \arctan\left(\frac{d_y}{d_x}\right) \tag{3.53}$$

Due to the absence of a mass term [80] ($d_3 = 0$), the angle ϕ is locked at $\pi/2$ and the versor field $\hat{\boldsymbol{d}}$ lies in the xy-plane.

The eigenstates of Eq. 3.50 are

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ +e^{i\theta} \end{pmatrix} \qquad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -e^{i\theta} \end{pmatrix} \tag{3.54}$$

with eigenvalues

$$\epsilon_{\mathbf{k}c} = +t|\Phi(\mathbf{k})| \qquad \epsilon_{\mathbf{k}v} = -t|\Phi(\mathbf{k})| \tag{3.55}$$

The stationary states are a combination of Bloch states associated to A and B sites (with the same Bloch vector $\mathbf{k}$), where the electron has an equal probability of being found in the sublattice A or B.

The non-abelian Berry connection of monolayer graphene is derived from the eigenstates (Eq. 3.26),

$$\mathcal{A}^{\alpha}_{\mathbf{k}cc} = \mathcal{A}^{\alpha}_{\mathbf{k}vv} = -\frac{1}{2}\,\partial^{\alpha}\theta \qquad\qquad \mathcal{A}^{\alpha}_{\mathbf{k}cv} = \mathcal{A}^{\alpha}_{\mathbf{k}vc} = \frac{1}{2}\,\partial^{\alpha}\theta \tag{3.56}$$

The dispersion relation is represented in Fig. 3.2. In undoped graphene, there is an electron per atom, which, taking into account the spin degeneracy, leads to the lower band being completely filled at zero temperature. The Fermi level is then at zero energy:

$$\Phi(\mathbf{k}) = 0 \tag{3.57}$$

The solutions for this last equation are $\mathbf{k} = \mathbf{K}$ or $\mathbf{k} = \mathbf{K}'$. These are the Dirac points. We will see why.

3.3.3 The Dirac Hamiltonian

The Dirac points constitute the Fermi surface of neutral graphene. If we examine the dispersion relation around these points, we find a rather unusual conical dispersion law, depicted in Fig. 3.2. Since the energy scales linearly with momentum, we can already infer that the excitations will be massless.

To analyze these excitations more quantitatively, let us expand $\Phi(\mathbf{k})$ around the Dirac point $\mathbf{K}$,

$$\begin{aligned}
\Phi(\mathbf{K}+\mathbf{q}) &\simeq \Phi(\mathbf{K}) + \mathbf{q}\cdot(\boldsymbol{\nabla}\Phi(\mathbf{k}))_{\mathbf{k}=\mathbf{K}} \\
&= i\mathbf{q}\cdot(e^{i\mathbf{K}\cdot\boldsymbol{\delta}_1}\,\boldsymbol{\delta}_1 + e^{i\mathbf{K}\cdot\boldsymbol{\delta}_2}\,\boldsymbol{\delta}_2 + e^{i\mathbf{K}\cdot\boldsymbol{\delta}_3}\,\boldsymbol{\delta}_3) \\
&= i\mathbf{q}\cdot(e^{i\frac{2\pi}{3}}\,\boldsymbol{\delta}_1 + e^{-i\frac{2\pi}{3}}\,\boldsymbol{\delta}_2 + \boldsymbol{\delta}_3) \\
&= ia\mathbf{q}\cdot\left(i\sin\left(\frac{2\pi}{3}\right), \frac{1}{\sqrt{3}}\left(-\cos\left(\frac{2\pi}{3}\right)+1\right)\right) \\
&= -\frac{\sqrt{3}a}{2}(q_x - iq_y)
\end{aligned} \tag{3.58}$$

and build a low energy effective Hamiltonian

$$H_0(\mathbf{q}) = -\frac{\sqrt{3}at}{2}\begin{pmatrix} 0 & q_x - i\,q_y \\ q_x + i\,q_y & 0 \end{pmatrix} \tag{3.59}$$

which can be succinctly expressed as a function of the momentum (relative to the Dirac point) $\mathbf{p} = \hbar\mathbf{q}$,

$$H_0 = v_F\,(\sigma_x\,p_x + \sigma_y\,p_y) \tag{3.60}$$

where $v_F = -3a_0t/2\hbar$ is the Fermi velocity. Introducing the parameters for graphene ($t \approx -3\ eV$), we find $v_F \approx 10^6\ m\,s^{-1}$.

Equation 3.60 is the Dirac equation for massless fermions. Since deviations from this behaviour are at energy scales of order 1 eV, we can say that electrons in graphene indeed act as massless Dirac fermions [66], when dealing with low energy excitations.

However, there is one relevant distinction from the analogous high-energy physics: σ is not the electron spin. It is an analogous variable, called pseudospin, where

$\sigma_z = +1$ for an electron in the sublattice A and $\sigma_z = -1$ for an electron in sublattice B. For the energy eigenstates, $< \sigma_z >= 1/2$.

In this low energy effective theory, we have $\boldsymbol{d}(\mathbf{q}) = \hbar v_F \, \mathbf{q}$. The dispersion relation is linear and gapless,

$$\epsilon_{\mathbf{q}c} = -\epsilon_{\mathbf{q}v} = \hbar v_F \, |\mathbf{q}| \tag{3.61}$$

The eigenstates still have the form in Eq. 3.54, but now the angle θ obeys

$$\theta = \arctan(q_y/q_x) \tag{3.62}$$

which implies that the pseudospin and momentum are locked in the same or opposite directions, for conduction and valence states, respectively.

The non-abelian Berry connection is now

$$\boldsymbol{\mathcal{A}}_{\mathbf{q}cv} = \boldsymbol{\mathcal{A}}_{\mathbf{q}vc} = -\boldsymbol{\mathcal{A}}_{\mathbf{q}cc} = -\boldsymbol{\mathcal{A}}_{\mathbf{q}vv} \tag{3.63}$$

$$\boldsymbol{\mathcal{A}}_{\mathbf{q}cv} = \frac{1}{2} \nabla\theta = \frac{1}{2\,q^2}(-q_y, q_x) = \frac{1}{2\,q}(-\sin\theta, \cos\theta) \tag{3.64}$$

Finally, the Berry curvature (Eq. 3.33) is zero everywhere, except for $\mathbf{q} = \mathbf{0}$. This is actually a consequence of graphene having both inversion and time-reversal symmetry (Section 6.6). At the Dirac points, the system is degenerate and the curvature is undefined.

We can get similar results for the other Dirac point (valley) $\mathbf{K}' = -\mathbf{K}$, simply by reflecting about the k_y axis:

$$H_0 = v_F \, (-\sigma_x \, p_x + \sigma_y \, p_y) \tag{3.65}$$

3.4 Gapped graphene or hBN

An important aspect of the previous two-band model to have in mind is that it respects inversion symmetry and therefore all even orders of the nonlinear conductivity will be automatically zero. The honeycomb lattice is not a Bravais lattice, but will be invariant under spatial inversion about the midpoint between an A and an B-site (of the same unit cell), as long as the atoms that occupy these sites are equivalent.

To enable a future study of the second order conductivity, an additional term is added to the previous graphene Hamiltonian that breaks the inversion symmetry of the crystal lattice and opens a gap in the density of states. Theoretically, this is done simply by considering the A and B atoms to have distinct on-site energies. Experimentally, gapped graphene can be achieved by means of a substrate induced-gap [81]. In such cases, the gap is usually small ($\Delta \approx 0.1 \; eV$).

Inversion symmetry can also be broken by placing two different atoms in the A and B sublattices. This is the case of hexagonal boron nitride, a wide gap two-dimensional insulator. Hexagonal boron nitride has attracted much attention as an excellent substrate for monolayer graphene, enhancing the electron mobilities by an order of magnitude [82], as well as for its role in more complex devices where it is combined with graphene in the so-called van der Walls heterostructures [63].

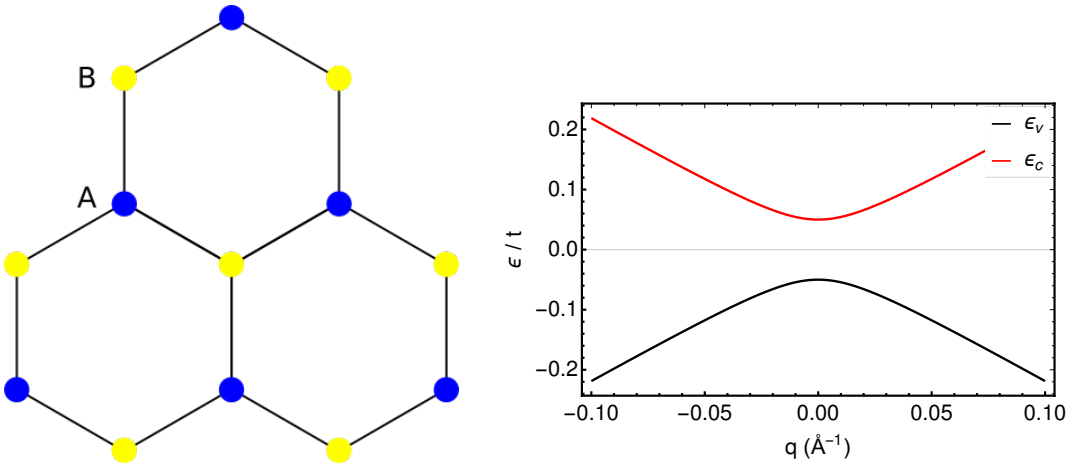


Figure 3.3: In the left, the crystal structure of hexagonal boron nitride. Boron atoms (blue) occupy the A-sites and nitrogen atoms (yellow) take the B-sites. In the right, the dispersion near the Dirac point for a gap $\Delta = 0.1\,t$.

3.4.1 Tight binding model

The tight binding model is similar to that of graphene, with identical hoppings, but different energies for A- and B-site orbitals (see Fig. 3.3),

$$\begin{aligned}\hat{H}_0 =& t\sum_{\mathbf{R}_n} |\varphi_A(\mathbf{r}-\mathbf{R}_n)\rangle \left(\langle\varphi_B(\mathbf{r}-\mathbf{R}_n)| + \langle\varphi_B(\mathbf{r}-\mathbf{R}_n-\mathbf{a}_1)| + \langle\varphi_B(\mathbf{r}-\mathbf{R}_n-\mathbf{a}_2)|\right) + h.c. \\ &+ \frac{\Delta}{2}\sum_{\mathbf{R}_n} |\varphi_A(\mathbf{r}-\mathbf{R}_n)\rangle\langle\varphi_A(\mathbf{r}-\mathbf{R}_n)| - \frac{\Delta}{2}\sum_{\mathbf{R}_n} |\varphi_B(\mathbf{r}-\mathbf{R}_n)\rangle\langle\varphi_B(\mathbf{r}-\mathbf{R}_n)| \end{aligned} \tag{3.66}$$

Following the same procedure as before, the Hamiltonian is represented in the sublattice Bloch basis,

$$H_0 = \begin{pmatrix} \Delta/2 & t\,\Phi(\mathbf{k}) \\ t\,\Phi^*(\mathbf{k}) & -\Delta/2 \end{pmatrix} \tag{3.67}$$

We still have $d_x = t\ \mathrm{Re}\{\Phi(\mathbf{k})\}$ and $d_y = -t\ \mathrm{Im}\{\Phi(\mathbf{k})\}$, but the added mass term $d_z = \Delta/2$ implies that $\boldsymbol{d}$ is no longer confined to a plane,

$$\phi = \arccos\left(\frac{\Delta}{\sqrt{4\,t^2\,|\Phi(\mathbf{k})|^2 + \Delta^2}}\right) \qquad \theta = \arctan\left(\frac{d_y}{d_x}\right) \tag{3.68}$$

There are also no longer any crossing points in the band structure,

$$\epsilon_{\mathbf{k}c} = +\sqrt{t^2\,|\Phi(\mathbf{k})|^2 + (\Delta/2)^2} \qquad \epsilon_{\mathbf{k}v} = -\sqrt{t^2\,|\Phi(\mathbf{k})|^2 + (\Delta/2)^2} \tag{3.69}$$

At the Dirac points, $\epsilon_{\mathbf{K}} = \epsilon_{\mathbf{K}'} = \pm\Delta/2$. The mass term lifted the degeneracy and introduced a gap of size Δ. This is represented in Fig. 3.3.

The energy eigenstates,

$$\begin{pmatrix} \cos(\phi/2) \\ \sin(\phi/2)e^{i\theta} \end{pmatrix} \qquad \begin{pmatrix} \sin(\phi/2) \\ -\cos(\phi/2)e^{i\theta} \end{pmatrix} \tag{3.70}$$

and the Berry connection can be found with the use of previously presented formulas (Eqs. 3.25 and 3.26).

If the Fermi surface lies in the gap or near the top (bottom) of the valence (conduction) band and for excitations of not too high an energy, the optical response will, at least in part, be determined by the electronic properties near the Dirac points. Once again, a low energy effective theory can be formulated around these points.

3.4.2 Low energy effective Hamiltonian

The expansion $\mathbf{k} \approx \mathbf{K} + \mathbf{q}$ is exactly the same as before and results in the following Hamiltonian,

$$H_0(\mathbf{q}) = \begin{pmatrix} \Delta/2 & \hbar v_F(q_x - i\, q_y) \\ \hbar v_F(q_x + i\, q_y) & -\Delta/2 \end{pmatrix} \tag{3.71}$$

which can be more concisely stated

$$H_0 = v_F\left(\sigma_x\, p_x + \sigma_y\, p_y\right) + \frac{\Delta}{2}\,\sigma_z \tag{3.72}$$

with $\mathbf{p} = \hbar\mathbf{q}$. Similar results can be derived for the other valley at $\mathbf{K}'$.

For this effective Hamiltonian, $\boldsymbol{d}(\mathbf{q}) = (\hbar v_F\, q_x, \hbar v_F\, q_y, \Delta/2)$. The eigenvalues give the band structure,

$$\epsilon_{\mathbf{q}c} = +\sqrt{(\hbar q)^2\, v_F^2 + (\Delta/2)^2} \qquad \epsilon_{\mathbf{q}v} = -\sqrt{(\hbar q)^2\, v_F^2 + (\Delta/2)^2} \tag{3.73}$$

This is the dispersion relation for massive relativistic particles, with the Fermi velocity v_F again replacing the speed of light and a mass $m = \Delta/(2v_F^2)$.

The stationary states and the Berry connection will depend only on the direction of the $\boldsymbol{d}$ field. The notation can therefore be abbreviated by factoring $\hbar v_F$, $\boldsymbol{d}(\mathbf{q}) = \hbar v_F(q_x, q_y, \Delta/2\hbar v_F)$, and defining $\Delta' \equiv \Delta/\hbar\, v_F$.

The energy eigenstates are the same as in Eq. 3.70, but now the angles obey

$$\phi = \arccos\left(\frac{\Delta'}{\sqrt{4\,q^2 + \Delta'^2}}\right) \qquad \theta = \arctan\left(\frac{q_y}{q_x}\right) \tag{3.74}$$

The pseudospin is off the xy plane for $\Delta \neq 0$ and hence misaligned with the direction of the in-plane momentum.

From Eqs. 3.74, the abelian Berry connection is derived (Eq. 3.26),

$$\mathcal{A}^x_{\mathbf{q}cc} = \left(1 - \frac{\Delta'}{\sqrt{4\,q^2 + \Delta'^2}}\right)\frac{\sin\theta}{2\,q} \qquad \mathcal{A}^y_{\mathbf{q}cc} = -\left(1 - \frac{\Delta'}{\sqrt{4\,q^2 + \Delta'^2}}\right)\frac{\cos\theta}{2\,q} \tag{3.75}$$

$$\mathcal{A}^x_{\mathbf{q}vv} = \left(1 + \frac{\Delta'}{\sqrt{4\,q^2 + \Delta'^2}}\right)\frac{\sin\theta}{2\,q} \qquad \mathcal{A}^y_{\mathbf{q}vv} = -\left(1 + \frac{\Delta'}{\sqrt{4\,q^2 + \Delta'^2}}\right)\frac{\cos\theta}{2\,q} \tag{3.76}$$

The non-abelian Berry connection contains additional off-diagonal matrix elements that, since the matrix is Hermitian $\boldsymbol{\mathcal{A}}^*_{\mathbf{q}vc} = \boldsymbol{\mathcal{A}}_{\mathbf{q}cv}$, are specified completely by

$$\mathcal{A}^x_{\mathbf{q}cv} = -\frac{\sin\theta}{2\sqrt{q^2+(\Delta'/2)^2}} + \frac{i\,\Delta'\cos\theta}{4\,(q^2+(\Delta'/2)^2)} \tag{3.77}$$

$$\mathcal{A}^y_{\mathbf{q}cv} = +\frac{\cos\theta}{2\sqrt{q^2+(\Delta'/2)^2}} + \frac{i\,\Delta'\sin\theta}{4\,(q^2+(\Delta'/2)^2)} \tag{3.78}$$

Since the mass term breaks inversion symmetry, the Berry curvature no longer vanishes. In vector form, it points out of the plane

$$\Omega^z_c = \mathcal{F}^{xy}_c = -\frac{\Delta'}{4\,(q^2+(\Delta'/2)^2)^{3/2}} \tag{3.79}$$

With the degeneracy removed, the curvature is now also defined at the origin, $q = 0$, where it reaches its maximum value.

The Berry curvature is relevant to linear response theory due to the appearance of a term in the optical conductivity, the so-called anomalous Hall conductivity, involving the flow of curvature over the entire FBZ (Chern number). This will be discussed later, in Chapter 5. Sadly, when integrated over the entire FBZ, the curvature flow returns zero for hexagonal boron nitride and gapped graphene. This is because, while broken inversion symmetry allows a non-zero curvature, the curvature is odd in $\mathbf{k}$ as a consequence of time-reversal symmetry (Section 6.6).

Chapter 4

Minimal coupling method

In this chapter, the difficulties inherent to perturbative calculations in the velocity gauge are addressed. Following closely the treatment in [56], it is shown how calculations of nonlinear conductivities in the independent particle approximation can be done to any order and for any finite band model. The advantages of adopting a minimal coupling Hamiltonian are described. As an illustration, the nonlinear optical conductivity of monolayer graphene is calculated numerically.

4.1 Minimal coupling in a finite band model

As it currently stands, the minimal coupling, or velocity gauge, Hamiltonian is of little use in calculations of the nonlinear conductivity. For any realistic calculation, a finite number of bands is employed and the expressions derived from perturbation theory (Eqs. 2.72, 2.74 and 2.75) give unphysical answers, divergent in the DC limit.

This is somewhat puzzling, considering that the theory is equivalent, by means of an unitary transformation (Appendix A), to the length gauge formulation, which has no issues in providing sensible answers for finite band models derived from tight binding or density functional theory. It was argued in Section 2.4.3 that gauge invariance will inevitably be broken by a band truncation that violates sum rules. Still, this does not properly answer why such approximations are more easily dealt with in the length gauge framework. More precisely, should it not be possible to simply define a model with a finite number of bands, assume that whatever other bands exist would be of such character that they would not contribute[1], and obtain proper physical predictions? If the velocity gauge of Section 2.4.2 is used, unreasonable infrared divergences make it clear that the answer is no.

In actuality, there is a very fundamental difference between the two gauges, that was not properly appreciated until recently [56], and it concerns the form of the perturbation.

In the length gauge, the perturbation has always the form $\hat{\mathcal{V}}(t) = e\,\hat{\mathbf{r}} \cdot \boldsymbol{E}(t)$. It does not matter if $\hat{H}_0$ represents the crystalline potential as in Eq. 2.35 or if it is defined in the context of a tight binding model with a finite number of bands, the length gauge perturbation, and therefore the structure of the perturbation theory, is unaffected. As long as the eigenstates of $\hat{H}_0$ are Bloch states and the eigenvalues

[1] By being too far off in energy and/or having very small matrix elements for the non-abelian Berry connection or whatever characteristic is required for band truncation to be a reasonable approximation in the respective perturbative treatment.

provide the band structure, the length gauge treatment in Section 2.4.1 remains valid.

To make it completely general, the unperturbed Hamiltonian can be defined by

$$\hat{H}_0 = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a |\psi_{\mathbf{k}a}\rangle\, \epsilon_{\mathbf{k}a}\, \langle\psi_{\mathbf{k}a}| \tag{4.1}$$

which can have any number of bands. It encompasses both the possibility of a Schrödinger Hamiltonian (Eq. 2.35) with an infinite Hilbert subspace at each $\mathbf{k}$, and that of an arbitrary finite band model, derived, for instance, from a tight binding description (Eqs. 3.6 and 3.7). Once this Hamiltonian is specified, the formulas in Eqs. 2.61, 2.63 and 2.64 provide the linear and nonlinear conductivities.

In sharp contrast, the minimal coupling Hamiltonian is defined as

$$\hat{H}(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\boldsymbol{A}(t)) \tag{4.2}$$

and relies on the expansion of $\hat{H}_0$ on the potential vector to define the perturbation. It was presumed before that the unperturbed Hamiltonian was

$$\hat{H}_0 = \frac{\hat{p}^2}{2\,m} + V(\hat{\mathbf{r}}) \tag{4.3}$$

and the perturbation was derived from there (Eq. 2.67), with the consequent perturbation theory of Section 2.4.2 built on this assumption. If a different $\hat{H}_0$ were chosen, the perturbation theory would have resulted different. It is therefore not that surprising that contradictions are found and sum rules broken upon replacement of $\hat{H}_0$ by a finite band model. In other words, the essential difficulty is that, in a minimal coupling formulation, *the perturbation depends explicitly on* $\hat{H}_0$, unlike the length gauge.

Having identified the origin of the problem, the path to its resolution becomes clear. The perturbation theory of Section 2.4.2 can be reformulated by dropping any assumptions on the form of the unperturbed Hamiltonian, other than it has the periodicity of some Bravais lattice, so that Bloch's theorem applies and there is a well defined First Brillouin Zone (FBZ). In this case, it can be written as in Eq. 4.1 and contain any number of bands. From this starting point, a perturbative analysis will be developed that generalizes the previous velocity gauge treatment and is applicable to finite band models.

The first challenge is to find a way to implement minimal coupling for the Hamiltonian in Eq. 4.1. The usual procedure dictates that the substitution $\hat{\mathbf{p}} \to \hat{\mathbf{p}} + e\,\boldsymbol{A}(t)$ be made, but in this case $\hat{H}_0$ is not defined in terms of the position and momentum operators, but instead is expressed in terms of the band structure and respective Bloch states.

An alternative is suggested by a rewriting of Eq. 4.2, inspired by the unitary transformation that relates the length and velocity gauges,

$$\hat{H}(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\boldsymbol{A}(t)) = \hat{\mathcal{U}}(t)\, \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}})\, \hat{\mathcal{U}}^\dagger(t) \tag{4.4}$$

with $\hat{\mathcal{U}}(t) = e^{-i\,e\,\hat{\mathbf{r}}\cdot\boldsymbol{A}(t)/\hbar}$.

Minimal coupling can be described as performing a unitary transformation $\hat{\mathcal{U}}(t)$. This presents a natural way to generalize the approach by defining the unitary transformation in the (possibly) finite subspace of $\hat{H}_0$ in Eq. 4.1. This is certainly

possible, since the position operator is defined in any space generated by Bloch states as the covariant derivative (Section 2.3).

$$\hat{H} = e^{-i\,e\,\hat{\mathbf{r}}\cdot\boldsymbol{A}(t)/\hbar}\,\hat{H}_0\,e^{i\,e\,\hat{\mathbf{r}}\cdot\boldsymbol{A}(t)/\hbar} = e^{e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar}\,\hat{H}_0\,e^{-e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar} \tag{4.5}$$

This is the minimal coupling Hamiltonian that will be used in this chapter and for which we will retrace the steps made in Chapter 2 and derive a new, more general, perturbation theory.

For that, it is first necessary to isolate the perturbation in Eq. 4.5. This is done with use of the Baker-Hausdorff lemma [83]: For any two operators, $\hat{A}$ and $\hat{B}$, the product $e^{\hat{A}}\,\hat{B}\,e^{-\hat{A}}$ can be expressed as a series of commutators:

$$e^{\hat{A}}\,\hat{B}\,e^{-\hat{A}} = \hat{B} + [\hat{A},\hat{B}] + \frac{1}{2}[\hat{A},[\hat{A},\hat{B}]] + \cdots = \sum_{n=0}^{+\infty}\frac{1}{n!}\,[\hat{A},[...[\hat{A},\hat{B}]]...] \tag{4.6}$$

where, in the sum, $\hat{B}$ is commuted with $\hat{A}$ an n number of times. The proof is straightforward and demands only the Taylor expansion of the exponential functions and appropriate grouping of the resulting terms.

Replacing $\hat{A}$ by $e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar$ and $\hat{B}$ by $\hat{H}_0$,

$$\begin{aligned}\hat{H} &= e^{e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar}\,\hat{H}_0\,e^{-e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar}\\ &= \sum_{n=0}^{+\infty}\frac{e^n}{n!\,\hbar^n}\,[\hat{D}^{\alpha_n},[...[\hat{D}^{\alpha_1},\hat{H}_0]]...]\,A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)\\ &= \hat{H}_0 + \hat{\mathcal{V}}(t)\end{aligned} \tag{4.7}$$

where we have identified the perturbation. It is represented in the eigenbasis of $\hat{H}_0$ as

$$\mathcal{V}_{\mathbf{k}ab}(t) = \sum_{n=1}^{+\infty}\frac{e^n}{n!}\,h^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab}\,A^{\alpha_1}(t)\ldots A^{\alpha_n}(t) \tag{4.8}$$

with

$$h^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab} \equiv \hbar^{-n}\,[D^{\alpha_n},[\ldots,[D^{\alpha_1},H_0]]\,...]_{\mathbf{k}ab} \tag{4.9}$$

The commutators $h^{\alpha_1\ldots\alpha_n}$ are the coefficients in the expansion of the Hamiltonian on the optical fields. In linear order, the coefficient is the unperturbed velocity, as presented in Eq. 2.55,

$$h^{\alpha}_{\mathbf{k}ab} = \hbar^{-1}\,[D^{\alpha},H_0]_{\mathbf{k}ab} = \frac{1}{\hbar}\,(\partial^{\alpha}\epsilon_{\mathbf{k}a})\,\delta_{ab} + \frac{i}{\hbar}\,\mathcal{A}^{\alpha}_{\mathbf{k}ab}\,\Delta\epsilon_{\mathbf{k}ab} \tag{4.10}$$

while for $n = 2$,

$$
\begin{aligned}
h^{\alpha_1\alpha_2}_{\mathbf{k}ab} &= \hbar^{-2}\,[D^{\alpha_2},[D^{\alpha_1},H_0]]_{\mathbf{k}ab} = \hbar^{-1}\,[D^{\alpha_2},h^{\alpha_1}]_{\mathbf{k}ab} \\
&= \frac{1}{\hbar}\left(\partial^{\alpha_2}h^{\alpha_1}_{\mathbf{k}ab}\right) - \frac{i}{\hbar}\sum_c (\mathcal{A}^{\alpha_2}_{\mathbf{k}ac}h^{\alpha_1}_{\mathbf{k}cb} - h^{\alpha_1}_{\mathbf{k}ac}\mathcal{A}^{\alpha_2}_{\mathbf{k}cb}) \\
&= \frac{1}{\hbar^2}\left(\partial^{\alpha_2}\partial^{\alpha_1}\epsilon_{\mathbf{k}a}\right)\delta_{ab} + \frac{i}{\hbar^2}\left(\partial^{\alpha_2}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\right)\Delta\epsilon_{\mathbf{k}ab} + \frac{i}{\hbar^2}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\left(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab}\right) \\
&\quad + \frac{i}{\hbar^2}\,\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\left(\partial^{\alpha_1}\Delta\epsilon_{\mathbf{k}ab}\right) + \frac{1}{\hbar^2}\sum_c \left(\mathcal{A}^{\alpha_2}_{\mathbf{k}ac}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}cb}\,\Delta\epsilon_{\mathbf{k}cb} - \mathcal{A}^{\alpha_1}_{\mathbf{k}ac}\,\mathcal{A}^{\alpha_2}_{\mathbf{k}cb}\,\Delta\epsilon_{\mathbf{k}ac}\right)
\end{aligned}
\tag{4.11}
$$

and so on.

Notice that the h coefficients and therefore the Hamiltonian are completely specified by the band structure $\epsilon_{\mathbf{k}a}$ and the non-abelian Berry connection $\mathcal{A}_{\mathbf{k}ab}$.

For the special case of the $\hat{H}_0$ in Eq. 4.3, discussed in Chapter 2, $h^{\alpha_1\alpha_2}$ is a constant, independent of $\mathbf{k}$, irrelevant for the dynamics of the system.

$$
h^{\alpha_1\alpha_2}_{\mathbf{k}ab} = -\hbar^{-2}\,[r^{\alpha_2},[r^{\alpha_1},H_0]]_{\mathbf{k}ab} = -\frac{i}{\hbar\,m}\,[r^{\alpha_2},p^{\alpha_1}]_{\mathbf{k}ab} = \frac{1}{m}\,\delta_{ab}\,\delta^{\alpha_1\alpha_2} \tag{4.12}
$$

Consequently, all other coefficients obtained by additional differentiation, return zero.

$$
h^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab} = 0 \qquad\qquad (n>2) \tag{4.13}
$$

The treatment of Section 2.4.2 is therefore a special case of the one presented here, that can be reobtained by setting the second order coefficient to a constant, the inverse mass, and all higher order coefficients to zero.

The perturbation in Eq. 4.8 is notably beyond the scope of the perturbative treatment in Chapter 2, which was based around a linear coupling with the classical field. It is then necessary to revisit and generalize the perturbation theory of Sections 2.1 and 2.2.

4.2 Revisiting perturbation theory

The Hamiltonian is not the only quantity to be expressed as a powers series in this formulation. The velocity is defined in terms of a commutator with the Hamiltonian and therefore

$$
\begin{aligned}
v^{\beta}_{\mathbf{k}ab} = \dot{r}^{\beta}_{\mathbf{k}ab} &= -\frac{i}{\hbar}\,[r^{\beta},H]_{\mathbf{k}ab} = \frac{1}{\hbar}\,[D^{\beta},H]_{\mathbf{k}ab} \\
&= \sum_{n=0}^{+\infty}\frac{e^n}{n!\,\hbar^{n+1}}\,[D^{\beta},[D^{\alpha_n},\ldots,[D^{\alpha_1},H_0]]]_{\mathbf{k}ab}\,A^{\alpha_1}(t)\ldots A^{\alpha_n}(t) \\
&= \sum_{n=0}^{+\infty}\frac{e^n}{n!}\,h^{\alpha_1\ldots\alpha_n\beta}_{\mathbf{k}ab}\,A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)
\end{aligned}
\tag{4.14}
$$

A similar situation was encountered before, in the velocity gauge treatment of Section 2.4.2. There, the coupling to the optical field was actually quadratic, instead

of linear, resulting in an additional term in the velocity (Eq. 2.70), which in turn resulted in an extra term for the linear conductivity (Eq. 2.72). Along the same lines, the couplings in the Hamiltonian discussed here involve higher powers of the optical fields and lead to additional terms in the velocity and more complicated expressions for the linear and nonlinear conductivities.

The velocity operator is, in this case, a function of the potential vector and explicitly time-dependent. The dynamics of the averaged current depends on the time evolution of both the velocity and the density operators,

$$J^{\beta}(t) = \mathrm{Tr}\left(\hat{J}^{\beta}(t)\,\hat{\rho}(t)\right) = -e\ \mathrm{Tr}\left(\hat{v}^{\beta}(t)\,\hat{\rho}(t)\right) = -e\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b} v^{\beta}_{\mathbf{k}ba}(t)\,\rho_{\mathbf{k}ab}(t) \tag{4.15}$$

The expansion of the current on the optical fields must then be done in the density matrix and the velocity matrix elements simultaneously.

In the absence of an external field, the current is

$$J^{\beta(0)} = -e\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b} v^{\beta(0)}_{\mathbf{k}ba}\,\rho^{(0)}_{\mathbf{k}ab} \tag{4.16}$$

The first order response is

$$J^{\beta(1)}(t) = -e\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(v^{\beta(1)}_{\mathbf{k}ba}(t)\,\rho^{(0)}_{\mathbf{k}ab} + v^{\beta(0)}_{\mathbf{k}ba}\,\rho^{(1)}_{\mathbf{k}ab}(t)\right) \tag{4.17}$$

Similarly for the second order response,

$$J^{\beta(2)}(t) = -e\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(v^{\beta(2)}_{\mathbf{k}ba}(t)\,\rho^{(0)}_{\mathbf{k}ab} + v^{\beta(1)}_{\mathbf{k}ba}(t)\,\rho^{(1)}_{\mathbf{k}ab}(t) + v^{\beta(0)}_{\mathbf{k}ba}\,\rho^{(2)}_{\mathbf{k}ab}(t)\right) \tag{4.18}$$

and, in general,

$$J^{\beta(n)}(t) = -e\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\sum_{p=0}^{n} v^{\beta(p)}_{\mathbf{k}ba}(t)\,\rho^{(n-p)}_{\mathbf{k}ab}(t) \tag{4.19}$$

The notation adopted here is the same as before. Particularly useful are the quantities $\rho^{\alpha_1\ldots\alpha_n}$ defined in Eqs. 2.20 and 2.22. In this formulation, the potential vector takes the role of the classical field the electron system couples to and the aforementioned definitions become

$$\hat{\rho}^{(n)}(t) \equiv \int_{-\infty}^{+\infty} dt_n \cdots \int_{-\infty}^{+\infty} dt_1\,\hat{\rho}^{\alpha_1\ldots\alpha_n}(t-t_1,\ldots,t-t_n)\,A^{\alpha_1}(t_1)\ldots A^{\alpha_n}(t_n) \tag{4.20}$$

which, in the frequency domain, translates to

$$\begin{aligned}\hat{\rho}^{(n)}(t) &\equiv \int_{-\infty}^{+\infty}\frac{d\omega_n}{2\,\pi}\ldots\int_{-\infty}^{+\infty}\frac{d\omega_1}{2\,\pi}\,\hat{\rho}^{\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n)\,A^{\alpha_1}(\omega_1)\ldots A^{\alpha_n}(\omega_n)\,e^{-i(\omega_1+\ldots+\omega_n)t}\\ &= \int_{-\infty}^{+\infty}\frac{d\omega_n}{2\,\pi}\ldots\int_{-\infty}^{+\infty}\frac{d\omega_1}{2\,\pi}\,\frac{(-i)^n}{\omega_1\ldots\omega_n}\,\hat{\rho}^{\alpha_1\ldots\alpha_n}(\omega_1,\ldots,\omega_n)\,E^{\alpha_1}(\omega_1)\ldots E^{\alpha_n}(\omega_n)\,e^{-i(\omega_1+\ldots+\omega_n)t}\end{aligned} \tag{4.21}$$

The previous expansion of the velocity (Eq. 4.14) can also be expressed in frequency space,

$$v^{\beta(n)}_{\mathbf{k}ab}(t) = \frac{e^n}{n!}\int_{-\infty}^{+\infty}\frac{d\omega_n}{2\,\pi}...\int_{-\infty}^{+\infty}\frac{d\omega_1}{2\,\pi}\,\frac{(-i)^n}{\omega_1...\omega_n}h^{\alpha_1...\alpha_n\beta}_{\mathbf{k}ab}\,E^{\alpha_1}(\omega_1)...E^{\alpha_n}(\omega_n)\,e^{-i(\omega_1+...+\omega_n)t} \tag{4.22}$$

Introducing Eqs. 4.21 and 4.22 in 4.19, we obtain formulas for the conductivities that are a generalization of 2.29. As in Section 2.5, their domain is then extended by considering complex frequencies. The linear conductivity is[2]

$$\sigma^{\beta\alpha}(\bar{\omega}) = \frac{i\,e}{\bar{\omega}}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(h^{\beta}_{\mathbf{k}ba}\,\rho^{\alpha}_{\mathbf{k}ab}(\bar{\omega}) + e\,h^{\alpha\beta}_{\mathbf{k}ba}\,(\rho_0)_{\mathbf{k}ab}\right) \tag{4.23}$$

The second order conductivity,

$$\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \frac{e}{\bar{\omega}_1\,\bar{\omega}_2}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(h^{\beta}_{\mathbf{k}ba}\,\rho^{\alpha_1\alpha_2}_{\mathbf{k}ab}(\bar{\omega}_1,\bar{\omega}_2) + e\,h^{\alpha_1\beta}_{\mathbf{k}ba}\,\rho^{\alpha_2}_{\mathbf{k}ab}(\bar{\omega}_2) + \frac{e^2}{2}\,h^{\alpha_1\alpha_2\beta}_{\mathbf{k}ba}\,(\rho_0)_{\mathbf{k}ab}\right) \tag{4.24}$$

The third order conductivity,

$$\begin{aligned}\sigma^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = -\frac{i\,e}{\bar{\omega}_1\,\bar{\omega}_2\,\bar{\omega}_3}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\Big(&h^{\beta}_{\mathbf{k}ba}\,\rho^{\alpha_1\alpha_2\alpha_3}_{\mathbf{k}ab}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3)\\ &+e\,h^{\alpha_1\beta}_{\mathbf{k}ba}\,\rho^{\alpha_2\alpha_3}_{\mathbf{k}ab}(\bar{\omega}_2,\bar{\omega}_3) + \frac{e^2}{2}\,h^{\alpha_1\alpha_2\beta}_{\mathbf{k}ba}\,\rho^{\alpha_3}_{\mathbf{k}ab}(\bar{\omega}_3) + \frac{e^3}{3!}\,h^{\alpha_1\alpha_2\alpha_3\beta}_{\mathbf{k}ba}\,(\rho_0)_{\mathbf{k}ab}\Big)\end{aligned} \tag{4.25}$$

Finally, the n-th order nonlinear conductivity can be written as

$$\sigma^{\beta\alpha_1...\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) = -e\sum_{p=0}^{n}\frac{(-i)^n}{\bar{\omega}_1\ldots\bar{\omega}_n}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{e^p}{p!}\,h^{\alpha_1...\alpha_p\beta}_{\mathbf{k}ba}\,\rho^{\alpha_{p+1}...\alpha_n}_{\mathbf{k}ab}(\bar{\omega}_{p+1},\ldots,\bar{\omega}_n) \tag{4.26}$$

The $\bar{\omega}_i^{-1}$ factors are, as in the previous velocity gauge treatment, due to the conversion of the potential vector into the electric field components $A(\omega) = -i\,E(\omega)/\omega$.

The functions $\rho^{\alpha_1...\alpha_n}$ still have to be evaluated. For this, the density matrix equation of motion must be solved,

$$i\,\partial_t\rho_{\mathbf{k}ab} = [H,\rho]_{\mathbf{k}ab} = [H_0,\rho]_{\mathbf{k}ab} + \sum_{n=1}^{+\infty}\frac{e^n}{n!}\,[h^{\alpha_1...\alpha_n},\rho(t)]_{\mathbf{k}ab}\;A^{\alpha_1}(t)\ldots A^{\alpha_n}(t) \tag{4.27}$$

A Fourier transform is applied to give

$$(\hbar\omega - \Delta\epsilon_{\mathbf{k}ab})\;\rho_{\mathbf{k}ab}(\omega) = \sum_{n=1}^{+\infty}\frac{e^n}{n!}\int_{-\infty}^{+\infty}[h^{\alpha_1...\alpha_n},\rho(t)]_{\mathbf{k}ab}\;A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)\,e^{i\omega t}\,dt \tag{4.28}$$

[2]This equation is a more general version of Eq. 2.72.

The resulting recursion relation is a bit more complicated than before (Eq. 2.14). At zero-th order, the equilibrium distribution is still the Fermi-Dirac distribution $\hat{\rho}^{(0)} = \hat{\rho}_0$.

In linear order,

$$\rho^{\alpha}_{\mathbf{k}ab}(\bar{\omega}) = \frac{e}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \left[h^{\alpha}, \rho_0\right]_{\mathbf{k}ab} \tag{4.29}$$

In second order,

$$\rho^{\alpha_1\alpha_2}_{\mathbf{k}ab}(\bar{\omega}_1, \bar{\omega}_2) = \frac{1}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab}} \left(e\left[h^{\alpha_1}, \rho^{\alpha_2}(\bar{\omega}_2)\right]_{\mathbf{k}ab} + \frac{e^2}{2}\left[h^{\alpha_1\alpha_2}, \rho_0\right]_{\mathbf{k}ab} \right) \tag{4.30}$$

The pattern is already becoming clear. As an additional example, the third order density matrix is provided by

$$\rho^{\alpha_1\alpha_2\alpha_3}_{\mathbf{k}ab}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \frac{1}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 + \hbar\bar{\omega}_3 - \Delta\epsilon_{\mathbf{k}ab}} \times \left(e\left[h^{\alpha_1}, \rho^{\alpha_2\alpha_3}(\bar{\omega}_2, \bar{\omega}_3)\right]_{\mathbf{k}ab} + \frac{e^2}{2}\left[h^{\alpha_1\alpha_2}, \rho^{\alpha_3}(\bar{\omega}_3)\right]_{\mathbf{k}ab} + \frac{e^3}{3!}\left[h^{\alpha_1\alpha_2\alpha_3}, \rho_0\right]_{\mathbf{k}ab} \right) \tag{4.31}$$

Finally, to general order n, the perturbative solution to the density matrix equation of motion is recursively expressed as

$$\rho^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab}(\bar{\omega}_1, \ldots, \bar{\omega}_n) = \frac{1}{\hbar\bar{\omega}_1 + \ldots + \hbar\bar{\omega}_n - \Delta\epsilon_{\mathbf{k}ab}} \sum_{m=1}^{n} \frac{e^m}{m!} \left[h^{\alpha_1\ldots\alpha_m}, \rho^{\alpha_{m+1}\ldots\alpha_n}(\bar{\omega}_{m+1}, \ldots, \bar{\omega}_n)\right]_{\mathbf{k}ab} \tag{4.32}$$

This recursion relation can be unfolded into lengthy expressions and its structure analyzed in more detail. However, we shall see that the real value of these expressions lies in their numerical evaluation (Section 4.4), for which a recursion relation is sufficient. Once the density matrix is computed via Eq. 4.32 it can inserted in Eq. 4.26 to give the nonlinear conductivity.

The nonlinear optical conductivity will still have to undergo the usual symmetrization procedure to ensure intrinsic permutation symmetry. Albeit trivial, this last step is a bit cumbersome to write down and will be left implicit. The new expressions for the conductivity presented here are entirely equivalent to the ones derived in Section 2.4.1 using the length gauge. Although far more complicated, they have their advantages, which will be discussed later.

4.3 Gauge invariance and sum rules

When, back in Chapter 2, two perturbation theories were developed to find the nonlinear conductivity, each employing a particular choice of gauge, but in principle completely equivalent, it was mentioned that their equivalence relied on sum rules. The length gauge expressions for the nonlinear conductivity (Eqs. 2.61, 2.63 and 2.64) are the same as the ones derived via the velocity gauge (Eqs. 2.72, 2.74 and 2.75), only if the following commutation relation holds

$$[\hat{r}^\beta, [\hat{r}^\alpha, \hat{H}_0]] = -\frac{\hbar^2}{m}\,\delta^{\beta\alpha} \tag{4.33}$$

which is true for the Hamiltonian in Eq. 4.3, but not necessarily for the one in Eq. 4.1.

The content of the previous sections generalizes the velocity gauge beyond such constraints. This condition is no longer necessary to demonstrate the equivalence of the recently derived expressions for the nonlinear conductivity (Eqs. 4.23, 4.24 and 4.25) and the length gauge ones (Eqs. 2.61, 2.63 and 2.64).

However, as it turns out, there are still some requirements for keeping the formulations equivalent or, more precisely, for maintaining the validity of the minimal coupling formulation: *the Hamiltonian* in Eq. 4.1 *must be defined over all the First Brillouin Zone* and the integration in Eq. 4.26 must run through its entirety. Low energy effective Hamiltonians that portray the electronic properties in a confined region of the FBZ are sometimes used (e.g. Eqs 3.60 and 3.72 for graphene) and often provide an accurate and simpler means for computing the optical response. For the minimal coupling approach presented in this chapter, however, these low energy descriptions do not suffice.

To understand why this is the case, we inspect the unitary transformation linking the length and velocity (minimal coupling) descriptions. The subscripts L and V are used to refer to the length and velocity gauges, respectively. The operators in the two descriptions are related by (Appendix A)

$$\hat{\mathcal{O}}_V = \hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L\,\hat{\mathcal{U}}^\dagger(t) \tag{4.34}$$

with

$$\hat{\mathcal{U}}(t) = e^{-i\,e\,\hat{\mathbf{r}}\cdot\boldsymbol{A}(t)/\hbar} \tag{4.35}$$

For a system with a (possibly) finite number of bands,

$$\hat{\mathcal{U}}(t) = e^{-i\,e\,\hat{\mathbf{r}}\cdot\boldsymbol{A}(t)/\hbar} = e^{e\,\hat{\mathbf{D}}\cdot\boldsymbol{A}(t)/\hbar} \tag{4.36}$$

Our quantities of interest are the ensemble average of the electric current and the linear and nonlinear conductivities. It seems, at first, straightforward to prove their equivalence,

$$J_L^\beta = \mathrm{Tr}\left(\hat{J}_L^\beta\,\hat{\rho}_L\right) = \mathrm{Tr}\left(\left(\hat{\mathcal{U}}(t)\,\hat{J}_L^\beta\,\hat{\mathcal{U}}^\dagger(t)\right)\left(\hat{\mathcal{U}}(t)\,\hat{\rho}_L\,\hat{\mathcal{U}}^\dagger(t)\right)\right) = \mathrm{Tr}\left(\hat{J}_V^\beta\,\hat{\rho}_V\right) = J_V^\beta \tag{4.37}$$

where it was made use of the cyclic property of the trace and the unitarity of the transformation.

This proof seems rather trivial, but since the position operator is the covariant derivative, the unitary transformation involves differentiation and the cyclic property of the trace is not quite sufficient to make the passage, being necessary also to perform an integration by parts, throwing away the integral of a gradient.

The cleanest way do demonstrate this is to write the transformation as a power series, in a manner entirely analogous to the minimal coupling Hamiltonian in Eq. 4.7,

$$
\begin{aligned}
J_V^\beta &= \mathrm{Tr}\left(\hat{\mathcal{U}}(t)\,\hat{J}_L^\beta\,\hat{\rho}_L\,\hat{\mathcal{U}}^\dagger(t)\right) \\
&= \sum_{n=0}^{+\infty} \frac{e^n}{n!\,\hbar^n}\, A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)\ \mathrm{Tr}\left([\hat{D}^{\alpha_n},[\ldots[\hat{D}^{\alpha_1},\hat{J}_L^\beta\,\hat{\rho}_L]]\ldots]\right) \\
&= J_L^\beta + \sum_{n=1}^{+\infty} \frac{e^n}{n!\,\hbar^n}\, A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)\ \mathrm{Tr}\left([\hat{D}^{\alpha_n},[\ldots[\hat{D}^{\alpha_1},\hat{J}_L^\beta\,\hat{\rho}_L]]\ldots]\right)
\end{aligned}
\tag{4.38}
$$

With this, we have arrived at what could be called a set of sum rules for the equivalence of the minimal coupling approach presented in this chapter and the length gauge method introduced in Section 2.4.1 (and further discussed in the next chapter):

$$
\sum_{n=1}^{+\infty} \frac{e^n}{n!\,\hbar^n}\, A^{\alpha_1}(t)\ldots A^{\alpha_n}(t)\ \mathrm{Tr}\left([\hat{D}^{\alpha_n},[\ldots[\hat{D}^{\alpha_1},\hat{J}_L^\beta\,\hat{\rho}_L]]\ldots]\right) = 0 \tag{4.39}
$$

These sum rules were first derived by Ventura *et al.* in [57].

The trace of a commutator is normally trivially zero, but since the covariant derivative is present, there is a contribution for each term in the sum that is proportional to

$$
\mathrm{Tr}\left([\hat{\partial}^{\alpha_n},[\ldots[\hat{D}^{\alpha_1},\hat{J}_L^\beta\,\hat{\rho}_L]\ldots]]\right) = \sum_a \int \frac{d^d\mathbf{k}}{(2\pi)^d}\left(\partial^{\alpha_n}[D^{\alpha_{n-1}},[\ldots[D^{\alpha_1},J_L^\beta\,\rho_L]\ldots]]_{\mathbf{k}aa}\right) = 0 \tag{4.40}
$$

This condition is satisfied due to the periodicity of the FBZ. But this implies all the FBZ must be used, otherwise the integration may not return zero and the length gauge and minimal coupling expressions will likely differ.

This argument feels somewhat abstract and is best understood by working the expressions in Eqs. 4.23, 4.24 and 4.25 into the format in Eqs. 2.61, 2.63 and 2.64, respectively, directly verifying the need for the sum rules in Eq. 4.40. Unfortunately, this type of manipulations are rather cumbersome. For this reason, they will be presented here only at first order. It should prove sufficient to make the previous argumentation concrete.

The goal is then to start with the expression for the linear conductivity obtained by a minimal coupling treatment (combining Eqs. 4.23 and 4.29),

$$
\begin{aligned}
\sigma^{\beta\alpha}(\bar{\omega}) &= \frac{i\,e}{\bar{\omega}} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \left(h^\beta_{\mathbf{k}ba}\,\rho^\alpha_{\mathbf{k}ab}(\bar{\omega}) + e\,h^{\alpha\beta}_{\mathbf{k}ba}\,(\rho_0)_{\mathbf{k}ab}\right) \\
&= \frac{i\,e^2}{\bar{\omega}} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \left(\frac{h^\beta_{\mathbf{k}ba}\,[h^\alpha,\rho_0]_{\mathbf{k}ab}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} + h^{\alpha\beta}_{\mathbf{k}ba}(\rho_0)_{\mathbf{k}ab}\right)
\end{aligned}
\tag{4.41}
$$

and arrive at the formula derived from the length gauge (rewriting Eq. 2.61 with the current notation),

$$
\sigma^{\beta\alpha}(\bar{\omega}) = -i\,e^2 \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{h^\beta_{\mathbf{k}ba}\,[D^\alpha,\rho_0]_{\mathbf{k}ab}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \tag{4.42}
$$

To begin, the Jacobi identity is used in a commutator from Eq. 4.41 to move the covariant derivative to the density matrix,

$$\begin{aligned}\hbar\left[h^{\alpha},\rho_0\right]_{\mathbf{k}ab} &= \left[\left[D^{\alpha},H_0\right],\rho_0\right]_{\mathbf{k}ab}\\ &= \left[\left[D^{\alpha},\rho_0\right],H_0\right]_{\mathbf{k}ab}+\left[D^{\alpha},\left[H_0,\rho_0\right]\right]_{\mathbf{k}ab}\\ &= \left[\left[D^{\alpha},\rho_0\right],H_0\right]_{\mathbf{k}ab}\end{aligned} \tag{4.43}$$

where in the last step we took into account that the commutator of two diagonal matrices is zero $[H_0,\rho_0]=0$.

This leads to

$$\left[h^{\alpha},\rho_0\right]_{\mathbf{k}ab} = -\hbar^{-1}\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}\Delta\epsilon_{\mathbf{k}ab} \tag{4.44}$$

With this, the first term in parenthesis of Eq. 4.41 becomes

$$\begin{aligned}\frac{h^{\beta}_{\mathbf{k}ba}\left[h^{\alpha},\rho_0\right]_{\mathbf{k}ab}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}} &= -\frac{h^{\beta}_{\mathbf{k}ba}\,\hbar^{-1}\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}\,\Delta\epsilon_{\mathbf{k}ab}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\\ &= h^{\beta}_{\mathbf{k}ba}\,\hbar^{-1}\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}-\frac{\bar{\omega}\,h^{\beta}_{\mathbf{k}ba}\,\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\end{aligned} \tag{4.45}$$

The second term in Eq. 4.45, when replaced in Eq. 4.41, will give the length gauge result in Eq. 4.42. The remaining contributions must therefore be zero and form our sum rule,

$$\frac{i\,e^2}{\bar{\omega}}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(h^{\beta}_{\mathbf{k}ba}\,\hbar^{-1}\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}+h^{\alpha\beta}_{\mathbf{k}ba}(\rho_0)_{\mathbf{k}ab}\right)=0 \tag{4.46}$$

This can be further simplified through

$$h^{\alpha\beta}_{\mathbf{k}ba}=\hbar^{-2}\left[D^{\beta},\left[D^{\alpha},H_0\right]\right]_{\mathbf{k}ba}=\hbar^{-2}\left[D^{\alpha},\left[D^{\beta},H_0\right]\right]_{\mathbf{k}ba}=\hbar^{-1}\left[D^{\alpha},h^{\beta}\right]_{\mathbf{k}ba} \tag{4.47}$$

where the commutation of covariant derivatives was used $\left[D^{\beta},D^{\alpha}\right]=0$ (see Appendix C). Replacing in Eq. 4.46,

$$\frac{i\,e^2}{\bar{\omega}}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(h^{\beta}_{\mathbf{k}ba}\,\hbar^{-1}\left[D^{\alpha},\rho_0\right]_{\mathbf{k}ab}+\hbar^{-1}\left[D^{\alpha},h^{\beta}\right]_{\mathbf{k}ba}(\rho_0)_{\mathbf{k}ab}\right)=0 \tag{4.48}$$

leading to

$$\frac{i\,e^2}{\hbar\bar{\omega}}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a}\left[D^{\alpha},h^{\beta}\rho_0\right]_{\mathbf{k}aa}=0 \tag{4.49}$$

The commutator with D^{α} can be broken into two pieces, one involving the Berry connection, which is trivially zero (the trace of a proper commutator is always zero) and another involving a conventional derivative,

$$\frac{i\,e^2}{\hbar\bar{\omega}}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(\partial^{\alpha}h^{\beta}_{\mathbf{k}ba}(\rho_0)_{\mathbf{k}ab}\right)=0 \tag{4.50}$$

which can be recognized as a particular case of the sum rules identified in Eq. 4.40.

This condition is always true, since the functions h and ρ_0 are periodic in reciprocal space. The sum rule (and therefore the equivalence between the results in the two gauges) is therefore trivially satisfied as long as the integral is performed over the full FBZ.

Analogous derivations can be made for higher orders in perturbation theory.

In the next chapter, it will be shown that in the length gauge the real part of the nonlinear conductivity is determined by specific regions of the FBZ, where resonance conditions are met. For this reason, it is not surprising that low energy effective theories can often be used in the length gauge to derive the optical response. In contrast, the minimal coupling formulation will not work with such models due to the previous argument: for the minimal coupling to provide the same, accurate, results as the length gauge, the sum rules in Eq. 4.40 must be met. Only by accident would such conditions be satisfied in a low energy effective model.

4.4 An efficient algorithm and its limitations

Let us assume then that we possess a model that is defined over a FBZ. The minimal coupling method is a new tool for computing the nonlinear conductivities, but how does it compare to the standard, and widely adopted, length gauge method? As usual, there are advantages and disadvantages associated with any particular choice of gauge. By considering the (exactly equivalent) forms of the nonlinear conductivities derived in the two gauges, the strengths and weaknesses of each can be analyzed.

A first look at Eq. 4.26 will immediately bring out the usual concerns with infrared divergences in the velocity gauge, due to all the inverse frequency factors. We emphasize again, however, that this expression is equivalent to the one obtained from the length gauge and therefore these divergences are only apparent. As it was exemplified in the previous section, the minimal coupling expressions can be manipulated and, using a series of sum rules, put in a form that is clearly divergence free in the DC limit. This approach was the one originally pursued [47] for the early velocity gauge calculations, but this use of sum rules became rather pointless after the length gauge formulation had been developed [49]. If the sum rules are employed in the velocity gauge to remove apparent divergences, one will simply arrive at an expression obtained more straightforwardly in the length gauge. With the current formulation of the velocity gauge, presented in this chapter, there is also no longer a risk of violating sum rules upon band truncation: Eq. 4.26 can be used without any worries of spurious infrared divergences.

Having clarified this point, it can still be noted that the minimal coupling version is considerably more elaborate; less useful not only for inspection, but in an actual analytical calculation. As an example, the expression of the conductivity responsible for second harmonic generation, with all components along the x axis, in the length gauge is,

$$\sigma^{xxx}(\bar{\omega},\bar{\omega}) = e^3 \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{h^x_{\mathbf{k}ba}}{2\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \left[D^x, \frac{1}{\hbar\bar{\omega} - \Delta\epsilon} \circ [D^x, \rho_0] \right]_{\mathbf{k}ab} \tag{4.51}$$

while in the minimal coupling formulation,

$$\sigma^{xxx}(\bar{\omega},\bar{\omega}) = \frac{e}{\bar{\omega}^2}\int \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(h^x_{\mathbf{k}ba}\,\rho^{xx}_{\mathbf{k}ab}(\bar{\omega},\bar{\omega}) + e\,h^{xx}_{\mathbf{k}ba}\,\rho^{x}_{\mathbf{k}ab}(\bar{\omega}) + \frac{e^2}{2}\,h^{xxx}_{\mathbf{k}ba}\,(\rho_0)_{\mathbf{k}ab}\right) \tag{4.52}$$

where we still have to write the density matrix components,

$$\begin{aligned}\rho^{x}_{\mathbf{k}ab}(\bar{\omega}) &= \frac{e\,[h^x,\rho_0]_{\mathbf{k}ab}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\\ \rho^{xx}_{\mathbf{k}ab}(\bar{\omega},\bar{\omega}) &= \frac{1}{2\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\left(e\,[h^x,\rho^x(\bar{\omega})]_{\mathbf{k}ab} + \frac{e^2}{2}\,[h^{xx},\rho_0]_{\mathbf{k}ab}\right)\end{aligned} \tag{4.53}$$

This example demonstrates that there is little advantage in doing the analytical calculations in the velocity gauge, although inspection of the previous equations shows an interesting point: there are only simple poles in the velocity gauge $(\hbar\omega - \Delta\epsilon)^{-1}$, while in the length gauge, by differentiation, higher order poles emerge. Still, for analytical calculations, I would advocate the cleaner and easier length gauge approach [49].

The strength of the minimal coupling method lies in the different arrangement of the commutators. The covariant derivatives are no longer applied to the density matrix in its recursion relation[3]. Instead, they operate only on the unperturbed Hamiltonian H_0 in the determination of the functions $h_{\mathbf{k}ab}$ (Eq. 4.9), which are independent of frequency, temperature and chemical potential.

These $h_{\mathbf{k}ab}$ functions are the essential objects in this formulation and the only that must be known analytically. The covariant derivative needs only to be applied to the unperturbed Hamiltonian, whereas in the length gauge it needs to be applied to the Fermi-Dirac distribution and the frequency poles, a more complex endeavor. Therein lies a significant advantage of the minimal coupling method.

A careful look at the algorithm delineated in Section 4.2, shows that for the nonlinear conductivity of order n, there are $n+1$ such functions to compute by successively applying a covariant derivative: $h^{\alpha_1\ldots\alpha_m}_{\mathbf{k}ab}$ with $m = 1, ..., n+1$. In the previous example of second harmonic generation, these would be $h^x_{\mathbf{k}ab}$, $h^{xx}_{\mathbf{k}ab}$ and $h^{xxx}_{\mathbf{k}ab}$. Further reducing this algorithm to its fundamental ingredients, we recognize once again that these calculations demand only a knowledge of two objects, which fully define the system under consideration: the dispersion relation $\epsilon_{\mathbf{k}a}$ and the Berry connection $\boldsymbol{\mathcal{A}}_{\mathbf{k}ab}$.

Once these $h^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab}$ functions are analytically determined, the integrand in Eq. 4.26 (or Eq. 4.52, in our example) can be numerically evaluated at each $\mathbf{k}$, independently and quite easily. In fact, the procedure involves evaluating the analytic $h^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab}$ functions and the Fermi-Dirac distribution at the $\mathbf{k}$ point and then computing simple commutators and traces of numeric matrices. There are no numerical derivatives at all. This is in contrast with the length gauge, where either the full expression of the response function is analytically calculated or numerical derivatives have to be applied in each step of the density matrix recursion relation. Either way, via the product rule and higher order poles, the number of complicated terms to evaluate grows very fast with n in the length gauge approach.

[3] In this aspect, the approach here has similarities with the one employed in [76].

For this reason, the form of the nonlinear conductivity in Eq. 4.26, derived from the minimal coupling Hamiltonian, should provide a more efficient numerical approach. The author has implemented numerically the expressions in both gauges and done calculations on the nonlinear conductivity of monolayer graphene and observed that the computation times were indeed considerably smaller when Eq. 4.26 was used.

Additionally, the frequency, temperature and chemical potential parameters can be changed at will ($h^{\alpha_1\cdots\alpha_n}$ is unchanged), without significant cost increase. This enables us to probe the response beyond what is usually possible to capture with low energy effective theories in the length gauge [84]. Also, considering perturbation theory beyond third order seems feasible, and straightforward, since it involves no substantial increase in the complexity of the calculation. This is hardly possible in a length gauge treatment.

The generalization of the velocity gauge discussed in this chapter was introduced in 2018 [56] and has since seen further developments by other researchers. Parker *et al.* [85] placed it in a diagrammatic form, discussed the low frequency limit where it meets the Boltzmann equation methods and used it to compute the nonlinear optical response of Weyl semimetals. S. João *et al* [86, 87] worked with a different diagrammatic formulation and developed the theory in an arbitrary basis. Numerical efficiency was achieved via the kernel polynomial method, making it possible to do the computations with very large numbers of atoms. By performing the calculations in real space, it became possible to introduce and study the effects of disorder, vacancies and lattice distortions.

4.5 Evaluating commutators for tight binding models

In this section, we tackle a subtle issue, albeit one that will prove of practical importance. As it was laid out in previous sections, numerical evaluation of the nonlinear conductivity via the minimal coupling method requires prior computation of the h commutators. Once these are known, we are left with numerical operations that can be done through a concise Mathematica code (or in some other appropriate programming language favored by the reader).

$h^{\alpha_1\cdots\alpha_n}$ is the result of successive covariant differentiation of the unperturbed Hamiltonian $\hat{H}_0$. This can be made the direct way: once the band structure and the non-abelian Berry connection are known, just expand the commutators in detail (as in Eqs. 4.10 and 4.11) and evaluate the derivatives of these quantities. There exists, however, a simpler way to compute these commutators for the tight-binding models used in this thesis. To understand this, it is necessary to discuss in a bit more depth the covariant derivative and reinterpret some formulas.

Traditionally, the parametric Hamiltonian is defined not as in Section 3.1, but by the following transformation,

$$\hat{H}_0(\mathbf{k}) = e^{-i\,\mathbf{k}\cdot\hat{\mathbf{r}}}\,\hat{H}_0\,e^{i\,\mathbf{k}\cdot\hat{\mathbf{r}}} \tag{4.54}$$

For the Hamiltonian in Eq. 4.3 and using $\hat{\mathbf{p}} = -i\hbar\,\nabla$,

$$\begin{aligned}\hat{H}_0(\mathbf{k}) &= e^{-i\,\mathbf{k}\cdot\hat{\mathbf{r}}}\left(-\frac{\hbar^2\,\nabla^2}{2\,m}+V(\hat{\mathbf{r}})\right)e^{i\,\mathbf{k}\cdot\hat{\mathbf{r}}}\\ &= \frac{\hbar^2}{2\,m}\left(\frac{\nabla}{i}+\mathbf{k}\right)^2+V(\hat{\mathbf{r}})\end{aligned} \tag{4.55}$$

$\hat{H}_0(\mathbf{k})$ represents a family of operators parametrized by the Bloch vector. The eigenvalues provide the band structure

$$\hat{H}_0(\mathbf{k})\,|u_{\mathbf{k}a}\rangle = \epsilon_{\mathbf{k}a}\,|u_{\mathbf{k}a}\rangle \tag{4.56}$$

and, by Bloch's theorem, the eigenstates are the periodic parts of the Bloch functions, from which the Berry connection is calculated,

$$\mathcal{A}^{\alpha}_{\mathbf{k}ab} \equiv i\,\langle u_{\mathbf{k}a}|\partial^{\alpha}u_{\mathbf{k}b}\rangle \tag{4.57}$$

This is the textbook presentation. It can be generalized by considering the unperturbed Hamiltonian to have the form of Eq. 4.1, where an arbitrary number of bands is considered. Then,

$$\hat{H}_0(\mathbf{k}) \equiv \sum_a |u_{\mathbf{k}a}\rangle\,\epsilon_{\mathbf{k}a}\,\langle u_{\mathbf{k}a}| \tag{4.58}$$

By the same reasoning, we can define for any operator diagonal in $\mathbf{k}$,

$$\hat{\mathcal{O}}(\mathbf{k}) \equiv \sum_{a,b} |u_{\mathbf{k}a}\rangle\,\mathcal{O}_{\mathbf{k}ab}\,\langle u_{\mathbf{k}b}| \tag{4.59}$$

This definition permit us to gather further insight on the covariant derivative. If the parametric derivative is taken in the previous equation and the matrix elements evaluated,

$$\begin{aligned}&\langle u_{\mathbf{k}a}|\,\partial^{\alpha}\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle = \Big(\partial^{\alpha}\,\langle u_{\mathbf{k}a}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle\Big) - \langle\partial^{\alpha}u_{\mathbf{k}a}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle - \langle u_{\mathbf{k}a}|\,\hat{\mathcal{O}}(\mathbf{k})\,|\partial^{\alpha}u_{\mathbf{k}b}\rangle\\ &= (\partial^{\alpha}\mathcal{O}_{\mathbf{k}ab}) + \sum_c\Big(-\langle\partial^{\alpha}u_{\mathbf{k}a}|u_{\mathbf{k}c}\rangle\,\langle u_{\mathbf{k}c}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle - \langle u_{\mathbf{k}a}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}c}\rangle\,\langle u_{\mathbf{k}c}|\partial^{\alpha}u_{\mathbf{k}b}\rangle\Big)\\ &= (\partial^{\alpha}\mathcal{O}_{\mathbf{k}ab}) + \sum_c\Big(+\langle u_{\mathbf{k}a}|\partial^{\alpha}u_{\mathbf{k}c}\rangle\,\langle u_{\mathbf{k}c}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle - \langle u_{\mathbf{k}a}|\,\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}c}\rangle\,\langle u_{\mathbf{k}c}|\partial^{\alpha}u_{\mathbf{k}b}\rangle\Big)\\ &= (\partial^{\alpha}\mathcal{O}_{\mathbf{k}ab}) - i\sum_c\left(\mathcal{A}^{\alpha}_{\mathbf{k}ac}\,\mathcal{O}_{\mathbf{k}cb} - \mathcal{O}_{\mathbf{k}ac}\,\mathcal{A}^{\alpha}_{\mathbf{k}cb}\right) = [D^{\alpha},\mathcal{O}]_{\mathbf{k}ab}\end{aligned} \tag{4.60}$$

We conclude that the operation of commuting with a covariant derivative can be interpreted as,

$$[D^{\alpha},\mathcal{O}]_{\mathbf{k}ab} = \langle u_{\mathbf{k}a}|\,\partial^{\alpha}\hat{\mathcal{O}}(\mathbf{k})\,|u_{\mathbf{k}b}\rangle \tag{4.61}$$

This is valid for any basis. If we define (Eq. 3.8),

$$\psi_{\mathbf{k}a}(\mathbf{r}) = \sum_{\lambda} c_{\mathbf{k}a\lambda}\,\Psi_{\mathbf{k}\lambda}(\mathbf{r}) \tag{4.62}$$

and

$$\Psi_{\mathbf{k}\lambda}(\mathbf{r}) = e^{i\,\mathbf{k}\cdot\mathbf{r}}\,\mathrm{u}_{\mathbf{k}\lambda}(\mathbf{r}) \tag{4.63}$$

where $\mathbf{u}_{\mathbf{k}\lambda}$ is the periodic part of the Bloch function, then

$$[D^\alpha, \mathcal{O}]_{\mathbf{k}\lambda\lambda'} = \partial^\alpha \mathcal{O}_{\mathbf{k}\lambda\lambda'} - i\,[A^\alpha, \mathcal{O}]_{\mathbf{k}\lambda\lambda'} = \langle \mathrm{u}_{\mathbf{k}\lambda}|\,\partial^\alpha \hat{\mathcal{O}}(\mathbf{k})\,|\mathrm{u}_{\mathbf{k}\lambda'}\rangle \tag{4.64}$$

with

$$A^\alpha_{\mathbf{k}\lambda\lambda'} \equiv i\,\langle \mathrm{u}_{\mathbf{k}\lambda}|\partial^\alpha \mathrm{u}_{\mathbf{k}\lambda'}\rangle \tag{4.65}$$

The quantity in Eq. 4.65 is, strictly speaking, not a Berry connection, since the $\mathrm{u}_{\mathbf{k}\lambda}$ do not correspond to the eigenstates of the Hamiltonian (Eq. 4.58).

Now, consider the case where the parametric Hamiltonian in Eq. 4.58 is derived from a tight binding model. Moreover, let us take $\Psi_{\mathbf{k}\lambda}$ to be the sublattice Bloch basis introduced in Eq. 3.4. In this case,

$$\hat{H}_0(\mathbf{k}) = \sum_{\lambda,\lambda'} |\mathrm{u}_{\mathbf{k}\lambda}\rangle\,(H_0)_{\mathbf{k}\lambda\lambda'}\,\langle \mathrm{u}_{\mathbf{k}\lambda'}| \tag{4.66}$$

with the matrix elements $(H_0)_{\mathbf{k}\lambda\lambda'}$ from Eq. 3.7.

The formula for the non-abelian Berry connection (Eq. 3.17) derived in Section 3.1 can be reinterpreted in light of the analysis of this section,

$$\begin{aligned}\mathcal{A}^\alpha_{\mathbf{k}ab} &= i\sum_\lambda c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda}) + \sum_{\lambda,\lambda'} c^*_{\mathbf{k}a\lambda}\,c_{\mathbf{k}b\lambda'}\,S^\alpha_{\mathbf{k}\lambda\lambda'} \\ &= i\sum_\lambda c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda}) + \sum_{\lambda} c^*_{\mathbf{k}a\lambda}\,c_{\mathbf{k}b\lambda}\,A^\alpha_{\mathbf{k}\lambda\lambda'}\end{aligned} \tag{4.67}$$

Tracing back the derivation of the second term in Eq. 3.17, it can be recognized that it is just the computation of the object in Eq. 4.65, $A = S$. The previously inspected limit of no overlap in the tight binding model can now be expressed as

$$A^\alpha_{\mathbf{k}\lambda\lambda'} = 0 \tag{4.68}$$

or

$$\mathcal{A}^\alpha_{\mathbf{k}ab} = i\sum_\lambda c^*_{\mathbf{k}a\lambda}(\partial^\alpha c_{\mathbf{k}b\lambda}) \tag{4.69}$$

Put succinctly, the computation of the non-abelian Berry connection is made by performing a change of basis from the stationary to the sublattice basis. Since the "Berry connection" A^α defined in the sublattice Bloch basis vanishes, the actual Berry connection $\mathcal{A}^\alpha$ can be obtained by working only with the coefficients $c_{\mathbf{k}a\lambda}$ that perform the change of basis. These coefficients are the eigenstates of the Hamiltonian representation in the sublattice Bloch basis $(H_0)_{\mathbf{k}\lambda\lambda'}$.

What is the significance of all this in finding the h commutators? In the absence of orbital overlap,

$$[D^\alpha, H_0]_{\mathbf{k}\lambda\lambda'} = \partial^\alpha (H_0)_{\mathbf{k}\lambda\lambda'} - i\,[A^\alpha, H_0]_{\mathbf{k}\lambda\lambda'} = \partial^\alpha (H_0)_{\mathbf{k}\lambda\lambda'} \tag{4.70}$$

For use in perturbation theory, the commutators must be expressed in the eigenbasis of $\hat{H}_0$. A change of basis leads to,

$$[D^{\alpha}, H_0]_{\mathbf{k}ab} = \sum_{\lambda,\lambda'} c^{*}_{\mathbf{k}a\lambda} \, [D^{\alpha}, H_0]_{\mathbf{k}\lambda\lambda'} \, c_{\mathbf{k}b\lambda'} = \sum_{\lambda,\lambda'} c^{*}_{\mathbf{k}a\lambda} \left(\partial^{\alpha}(H_0)_{\mathbf{k}\lambda\lambda'}\right) c_{\mathbf{k}b\lambda'} \tag{4.71}$$

Similarly,

$$[D^{\alpha_2}, [D^{\alpha_1}, H_0]]_{\mathbf{k}ab} = \sum_{\lambda,\lambda'} c^{*}_{\mathbf{k}a\lambda} \left(\partial^{\alpha_2}\partial^{\alpha_1}(H_0)_{\mathbf{k}\lambda\lambda'}\right) c_{\mathbf{k}b\lambda'} \tag{4.72}$$

and

$$[D^{\alpha_n}, \ldots [D^{\alpha_1}, H_0] \ldots]_{\mathbf{k}ab} = \sum_{\lambda,\lambda'} c^{*}_{\mathbf{k}a\lambda} \left(\partial^{\alpha_n}\ldots\partial^{\alpha_1}(H_0)_{\mathbf{k}\lambda\lambda'}\right) c_{\mathbf{k}b\lambda'} \tag{4.73}$$

We arrive then at the somewhat surprising conclusion that, for tight binding models with sufficiently localized orbitals, the process of evaluating the $h^{\alpha_1\cdots\alpha_n}$ commutators equates to differentiating the Hamiltonian in the sublattice Bloch basis and performing a mere change of basis. In particular, no knowledge of the non-abelian Berry connection is needed!

This interlude has then served its purpose, presenting a much simplified method to evaluate the h commutators, necessary ingredients in the minimal coupling method. There is not even need to express them analytically anymore: once the Hamiltonian H_0 is specified and differentiated, it can be evaluated at any given $\mathbf{k}$-point in the FBZ. The states and eigenvalues of the Hamiltonian are computed at that point by numerical diagonalization, the latter giving energies, the former the coefficients in Eq. 4.62. These are inserted in Eq. 4.73 to obtain the h commutators at the respective point in the FBZ.

A common concern is that, in the numerical diagonalization of the Hamiltonian, the phases of the eigenstates will not be defined continuously over the FBZ, but this is of no relevance here. The linear and the nonlinear conductivity, resulting from the evaluation of a trace, are invariant under the $U(1)$ gauge transformation of the Bloch functions[4].

The algorithm will be exemplified in the next section with a study of the nonlinear optical response of monolayer graphene.

4.6 Harmonic generation in monolayer graphene (numerical)

4.6.1 Setting up

In the final section of this chapter, the minimal coupling method is tested numerically, by computing the linear, second and third order conductivity of monolayer graphene. For conciseness, the analysis is restricted to the case of harmonic generation, where $\omega_1 = \omega_2 = \omega_3 = \omega$.

The basic physics of graphene and the nearest neighbour tight binding model that describes it were already introduced in Sections 3.3 and 3.4. From this description, the following Hamiltonian was constructed (Eq. 3.67),

[4] See discussion at the end of Section 2.3.

$$(H_0)_{\mathbf{k}\lambda\lambda'} = \begin{pmatrix} \Delta/2 & t\,\Phi(\mathbf{k}) \\ t\,\Phi^*(\mathbf{k}) & -\Delta/2 \end{pmatrix}_{\mathbf{k}\lambda\lambda'} \tag{4.74}$$

where the gap Δ was introduced by assigning different energies to the A- and B-site orbitals. The off-diagonal matrix elements are defined by (Eq. 3.49),

$$\begin{aligned}\Phi(\mathbf{k}) = &\cos\left(\frac{k_x\,a}{2} - \frac{k_y\,a}{2\sqrt{3}}\right) + \cos\left(-\frac{k_x\,a}{2} - \frac{k_y\,a}{2\sqrt{3}}\right) + \cos\left(\frac{k_y\,a}{\sqrt{3}}\right) \\ &+ i\,\sin\left(\frac{k_x\,a}{2} - \frac{k_y\,a}{2\sqrt{3}}\right) + i\,\sin\left(-\frac{k_x\,a}{2} - \frac{k_y\,a}{2\sqrt{3}}\right) + i\,\sin\left(\frac{k_y\,a}{\sqrt{3}}\right)\end{aligned} \tag{4.75}$$

Notice that this model is defined over the entire hexagonal FBZ of the honeycomb lattice (Fig. 3.1), a requirement for the minimal coupling method. This is unlike the low energy Hamiltonians in Eqs. 3.59 and 3.71, for which this method is not applicable.

The resulting band structure (Eqs. 3.69) is

$$\epsilon_{\mathbf{k}} = \pm\sqrt{\left(3 + 2\cos\left(k_x a\right) + 4\cos\left(\frac{k_x a}{2}\right)\cos\left(\frac{\sqrt{3}k_y a}{2}\right)\right)t^2 + \left(\frac{\Delta}{2}\right)^2} \tag{4.76}$$

Since this is a tight binding model, we might, under the aforementioned approximations (Section 4.5), dispense with the derivation of the Berry connection, and evaluate the $h^{\alpha_1\cdots\alpha_n}$ commutators by direct differentiation of the Hamiltonian.

Even though the prerequisites for the use of the minimal coupling algorithm are already satisfied and there is no need for knowing the $h^{\alpha_1\cdots\alpha_n}$ commutators analytically (indeed, the analytic form will not be used in the subsequent computations), it seems instructive to derive them at least once.

For this purpose, we momentarily consider the gapless case ($\Delta = 0$) and the response to an optical field aligned along the x axis. Using Eq. 4.73, with the Hamiltonian in Eq. 4.74 and the eigenstates in Eqs. 3.54,

$$h^{x}_{\mathbf{k}ab} = \frac{a\,\sin\left(\frac{k_x a}{2}\right)C_{ab}}{\hbar\sqrt{3 + 2\cos\left(k_x a\right) + 4\cos\left(\frac{k_x a}{2}\right)\cos\left(\frac{\sqrt{3}k_y a}{2}\right)}} \tag{4.77}$$

$$h^{xx}_{\mathbf{k}ab} = \frac{a^2\,\cos\left(\frac{k_x a}{2}\right)C_{ab}}{2\hbar^2\sqrt{3 + 2\cos\left(k_x a\right) + 4\cos\left(\frac{k_x a}{2}\right)\cos\left(\frac{\sqrt{3}k_y a}{2}\right)}} \tag{4.78}$$

$$h^{xxx}_{\mathbf{k}ab} = -\frac{a^2}{4\hbar^2}h^{x}_{\mathbf{k}ab} \qquad h^{xxxx}_{\mathbf{k}ab} = -\frac{a^2}{4\hbar^2}h^{xx}_{\mathbf{k}ab} \tag{4.79}$$

where C is a matrix in the band indices,

$$C_{ab} \equiv t\begin{pmatrix} 2\cos\left(\frac{k_x a}{2}\right) + \cos\left(\frac{\sqrt{3}k_y a}{2}\right) & -i\sin\left(\frac{\sqrt{3}k_y a}{2}\right) \\ i\sin\left(\frac{\sqrt{3}k_y a}{2}\right) & -2\cos\left(\frac{k_x a}{2}\right) + \cos\left(\frac{\sqrt{3}k_y a}{2}\right) \end{pmatrix}_{ab} \tag{4.80}$$

There is a cyclic character to the h commutators in this model. After the covariant derivative is applied more than twice, they repeat, aside from a numerical factor. This is due to the cosine and sine functions in Eq. 4.75.

It worth reminding that the commutators could have equally well been obtained by prior evaluation of the Berry connection (applying Eq. 4.69 to Eq. **??**) and expanding the commutators (Eqs. 4.10 and 4.11 and higher order analogues). This is a far more extensive procedure. Nonetheless, if we were to extend our analysis beyond tight binding models, it would be a necessary one.

The evaluation of the commutators in Eqs. 4.77-4.79 concludes all the setup necessary for a calculation of the nonlinear conductivity with the minimal coupling method, up to third order. However, attention should always be given to crystal symmetry, that may reduce the number of independent components and spare us unnecessary work.

A case in point, the symmetry of the honeycomb lattice [88] dictates that

$$\sigma^{xx}(\bar{\omega}) = \sigma^{yy}(\bar{\omega}) \qquad \sigma^{xy}(\bar{\omega}) = \sigma^{yx}(\bar{\omega}) = 0 \tag{4.81}$$

This is for the linear conductivity. The second order conductivity obeys

$$\sigma^{xxy}(\bar{\omega}_1, \bar{\omega}_2) = \sigma^{xyx}(\bar{\omega}_1, \bar{\omega}_2) = \sigma^{yxx}(\bar{\omega}_1, \bar{\omega}_2) = -\sigma^{yyy}(\bar{\omega}_1, \bar{\omega}_2) \tag{4.82}$$

$$\sigma^{yxx}(\bar{\omega}_1, \bar{\omega}_2) = \sigma^{yxy}(\bar{\omega}_1, \bar{\omega}_2) = \sigma^{xyy}(\bar{\omega}_1, \bar{\omega}_2) = \sigma^{xxx}(\bar{\omega}_1, \bar{\omega}_2) = 0 \tag{4.83}$$

If the gap is closed by taking $\Delta = 0$, then inversion symmetry is restored and the second order conductivity vanishes identically. Other than that, the tensor symmetries are the same, from linear to third order, for graphene with and without a band gap[5].

And in third order,

$$\sigma^{xxyy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \sigma^{yyxx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \tag{4.84}$$

$$\sigma^{xyxy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \sigma^{yxyx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \tag{4.85}$$

$$\sigma^{xyyx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \sigma^{yxxy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \tag{4.86}$$

$$\sigma^{xxxx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \sigma^{yyyy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \tag{4.87}$$

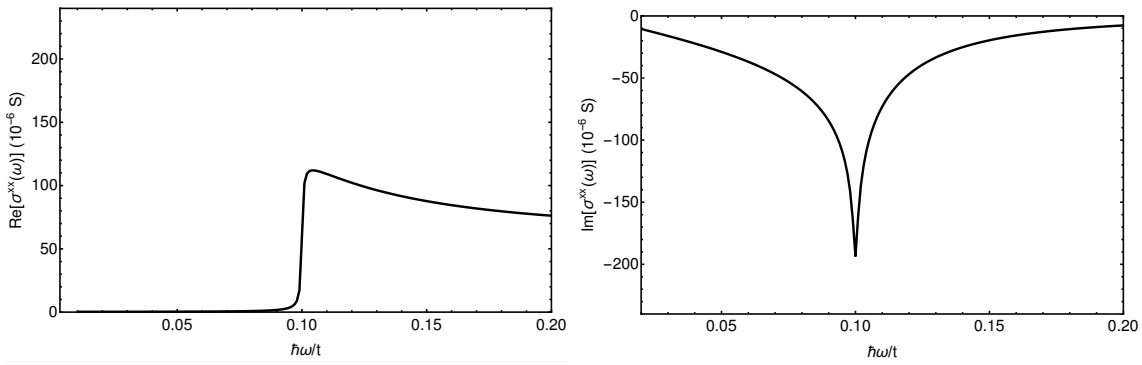
while the other, omitted, tensor components are all zero. There is one final identity, reducing the third order conductivity to three independent tensor elements:

$$\sigma^{xxxx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = \sigma^{xxyy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma^{xyxy}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma^{xyyx}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \tag{4.88}$$

In the special case of harmonic generation, intrinsic permutation symmetry implies

$$\sigma^{xxyy}(\bar{\omega}, \bar{\omega}, \bar{\omega}) = \sigma^{xyxy}(\bar{\omega}, \bar{\omega}, \bar{\omega}) = \sigma^{xyyx}(\bar{\omega}, \bar{\omega}, \bar{\omega}) = \sigma^{xxxx}(\bar{\omega}, \bar{\omega}, \bar{\omega})/3 \tag{4.89}$$

[5]The point group symmetry of graphene is D_{6h}, while that of gapped graphene or hexagonal boron nitride is D_{3d}. For the purposes of harmonic generation at linear and third order, the respective response functions are isotropic in both cases.



(a) (b)

Figure 4.1: Optical conductivity of gapped graphene with $\Delta = 0.1\,t$ and $\gamma = 5 \times 10^{-4}\,t$, at $T = 0\,\mathrm{K}$. The Fermi level resides in the band gap. The real part (a) has a step at $\omega = \Delta$, while the imaginary part (b) "diverges" to negative values at the same frequency. Since the system is an insulator, there is no Drude peak at zero frequency and both real and imaginary parts vanish for $\omega \to 0^+$.

With this, there is a single independent tensor component at each order. We need only to compute the components σ^{xx}, σ^{yyy} and σ^{xxxx} for the linear, second and third order optical response. These follow by feeding the band structure and the h commutators presented here into the algorithm, described in Section 4.4 and implemented in Mathematica code. The next sections display these linear and nonlinear conductivities.

4.6.2 Linear conductivity

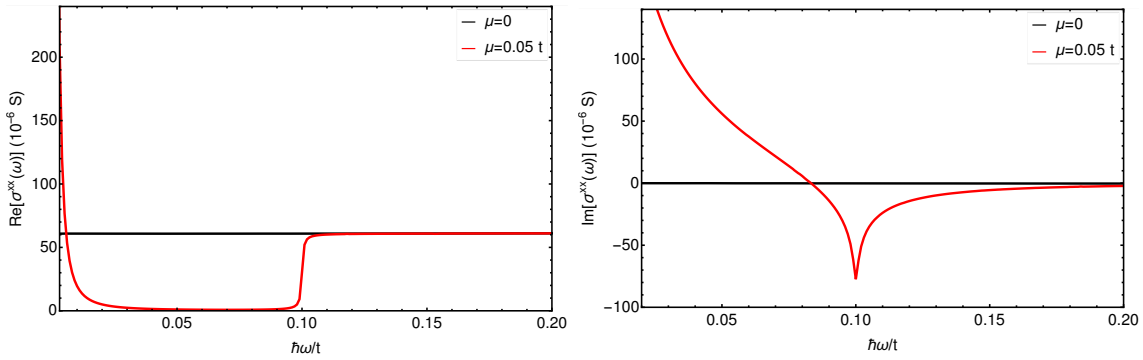
In this and the next subsections, the optical response of both gapped and gapless graphene is presented and discussed. The linear response is, naturally, the simplest and the first to be examined here. All the results are obtained at $T = 0\,\mathrm{K}$. Also, the existence of spin is accounted for by simply doubling the response, since it has no direct impact on the band structure or other electronic properties[6].

The optical conductivity of gapped graphene is depicted in Fig. 4.1. A modest value for the band gap was chosen: $\Delta = 0.1\,t \simeq 0.3$ eV, same as in Fig. 3.3. The Fermi level is placed in the gap.

It is worth noticing, considering the history with velocity gauge calculations (Section 2.4.3), that the DC conductivity of the gapped system is zero, as expected from an insulator. In Fig. 4.1, as in subsequent figures representing the nonlinear response of gapped graphene (Figs. 4.5 and 4.7), there are no unphysical infrared divergences, attesting to the validity of the minimal coupling method proposed in this chapter. This confirms that, despite appearances, Eq. 4.26 is well-behaved in the limit $\omega \to 0^+$.

In fact, the real part of the optical conductivity in Fig. 4.1a vanishes completely as soon as the optical frequency ω falls below the band gap. As will be proved in the next chapters, this a general characteristic for any time-reversal symmetric system without a Fermi surface. When the optical frequency matches the band gap, there

[6] Aside from affecting the counting of states, namely when relating the Fermi level to the carrier concentration.



(a) (b)

Figure 4.2: Optical conductivity of graphene ($\Delta = 0$) with $\gamma = 5 \times 10^{-4}\, t$, at $T = 0\,\mathrm{K}$. In black, the curves represent the case with no carriers, $\mu = 0$, and in red, the doped system with $\mu = 0.05\, t$, setting an effective gap at $2\,|\mu| = 0.1\, t$. For the undoped system, the real part (a) is defined by universal constants $\sigma_0 = \pi\, e^2/2\, h$ and it has no imaginary part (b). The response of the doped system is similar to Fig. 4.1; it differs in the existence of a Drude peak at zero frequency.

is a sudden, abrupt jump in the real part of the optical conductivity. Afterwards, the conductivity slowly decreases for higher frequencies, towards a constant value.

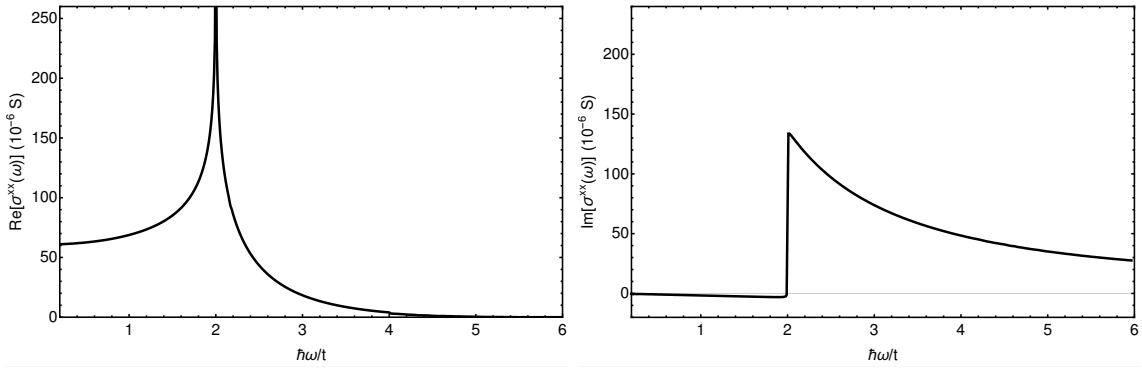
The step observed in the real part is matched by a sharp negative peak in the imaginary part, at the same frequency. This is, again, a common characteristic. It will be shown later, in Section 6.1, that there is a correspondence between the features in the real and imaginary parts, guaranteed by causality. The imaginary part of the response decays quickly in magnitude for higher frequencies and it is worth noting, again, that when $\omega \to 0^+$, it tends to zero.

When depicting the dispersion of the optical response of a crystal, the features will invariably be more pronounced at zero temperature and in the relaxation-free limit. This applies to both the linear and nonlinear conductivities. It is generally true that decreasing the parameter γ (Section 2.5) results in a cleaner, more physically transparent dispersion curve for the conductivity. The smaller the relaxation parameter, the better defined the features will be: the steeper the jumps in the real part and the sharper the peaks in the imaginary part. In particular, the peak in Fig. 4.1b falls deeper and deeper for smaller values of γ and ultimately diverges in the relaxation-free limit.

In this section, the relaxation is set at $\gamma = 0.0005\, t \simeq 1.5\,\mathrm{meV}$, which places it at $0.5\,\%$ of the next energy scale of the problem, the band gap. This leaves us comfortably in a regime where relaxation is (mostly) negligible.

The optical conductivity of graphene for $\Delta = 0$ is depicted in Fig. 4.2. When the Fermi level is set at zero energy, $\mu = 0$, crossing the Dirac points, the optical conductivity is purely real and constant over the entire frequency range represented in Fig. 4.2: $\sigma^{xx}(\omega) = \sigma_0 = e^2/4\,\hbar = \pi\, e^2/2\, h \simeq 60\,\mu\mathrm{S}$. This relates to a famous property of graphene, discovered early on [70]: a transparency (transmittance) in the visible defined by universal constants, $T_r = 1 - \pi\alpha_f$ with $\alpha_f = e^2/4\pi\epsilon_0\hbar c$ as the fine structure constant.

Another curious aspect of the linear conductivity of undoped graphene is that it remains constant for $\omega \to 0^+$. In other words, undoped graphene has a nonzero DC conductivity, even though there are no charge carriers! This formed the subject



(a) (b)

Figure 4.3: Optical conductivity of graphene ($\Delta = 0$) with $\mu = 0$ and $\gamma = 5 \times 10^{-4}\, t$, at $T = 0\,\mathrm{K}$. The optical response has a pronounced feature at $\hbar\omega = 2\,\epsilon_{\mathrm{vH}} = 2\,t$. The situation is inverted relative to the dispersion near the Dirac point (Fig. 4.2): the real part (a) displays a sharp peak, while the imaginary part (b) contains a sudden jump.

of much debate and the customary explanation for the observation of a metallic regime for very low carrier concentrations involves taking into account the existence of electron-hole puddles [89] and considering a more complex system than the clean, homogeneous crystal that concern us here. It is interesting though, that this astonishing property is quickly derived by perturbation theory. Unfortunately, the reasonableness of this perturbative result for $\omega = 0$ is highly questionable (see the discussion in Section 4.6.4).

The red curves in Fig. 4.2 depict the linear optical response of doped graphene. For $\mu \neq 0$, the Fermi level sets an effective gap at $2|\mu|$. Comparing with the gapped system in Fig. 4.1, we recognize the step in the real part once the photon frequency exceeds the effective gap and the corresponding drop in the imaginary part. This behavior is now due to Pauli blocking: at zero temperature, for energies below the Fermi level, both be valence and the conduction states are occupied and no vertical transitions are possible (see Fig. 4.4a).

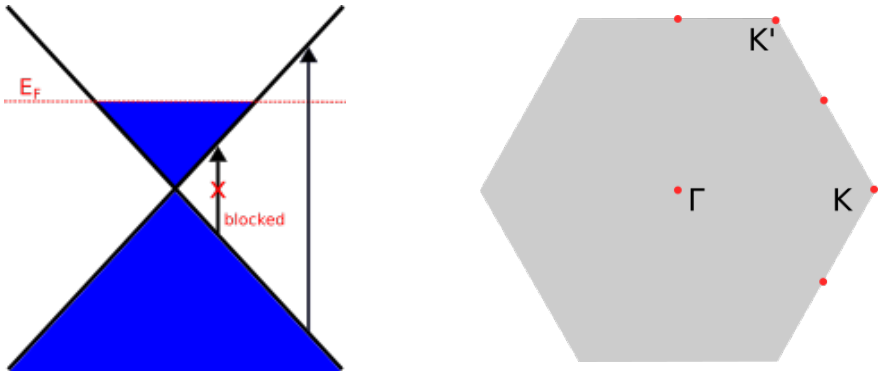
A distinct feature of doped graphene that separates it from the gapped system is the Drude peak at low frequencies. This contribution to the conductivity is determined by intraband transitions near the Fermi surface (more on this in Chapter 5). As it is to be expected from a metallic system with an concentration of free charge carriers, the conductivity tends to infinity[7] in the DC limit: this is a *physical* divergence.

So far, the results that have been presented are well-known in the literature, easily obtained by Kubo's formula. An advantage of the minimal coupling method adopted here is that we can effortlessly consider higher optical frequencies, probing regions of the FBZ beyond the Dirac cones. The dispersion of the optical conductivity of undoped graphene in an wider frequency range is represented in Fig. 4.3.

Interestingly, there is an additional, very prominent feature in the conductivity, larger in magnitude then the step and peak of Fig. 4.2 and localized around $\hbar\omega = 2\,t$. What is special about this specific frequency?

Besides combinations of optical frequencies that match the gap, unblocking in-

[7] For a finite relaxation parameter, it does not actually diverge, but is proportional to γ^{-1}.



(a) (b)

Figure 4.4: In the left, (a) a cross section of a Dirac cone, with arrows indicating electronic interband transitions caused by the incidence of a photon with that energy. For optical frequencies below the Fermi level, both the valence and conduction states are occupied and, by Pauli's exclusion principle, interband transitions are blocked. In the right, (b) van Hove singularities are marked in red in the FBZ of graphene. All other points where the energy gradient vanishes can be obtained by translation with a reciprocal lattice vector. $\mathbf{\Gamma}$ stands for the center of the FBZ, where the conduction band energy is maximum, $\mathbf{K}$ and $\mathbf{K}'$ are the Dirac points and the remaining van Hove singularities are responsible for the feature observed in Fig. 4.3.

terband transitions, another set of frequencies which have a considerable impact in the optical response are those for which the joint density of states diverges: the van Hove singularities [58].

To find these frequencies, one must first identify the points in the FBZ for which $\partial^\beta \Delta\epsilon_{\mathbf{k}cv} = 0$. Because graphene has electron hole symmetry, this simplifies to $\partial^\beta \Delta\epsilon_{\mathbf{k}cv} = 2\left(\partial^\beta \epsilon_{\mathbf{k}c}\right) = 0$. This energy will be denoted by ϵ_{vH},

$$|\nabla_{\mathbf{k}} \epsilon_{\mathrm{vH}}| = 0 \tag{4.90}$$

Inserting the dispersion relation (Eq. 4.76) in Eq. 4.90 leads into the following system of equations,

$$-2\,\sin(k_x\,a) - 2\,\sin\left(\frac{k_x\,a}{2}\right)\,\cos\left(\frac{\sqrt{3}\,k_y\,a}{2}\right) = 0 \tag{4.91}$$

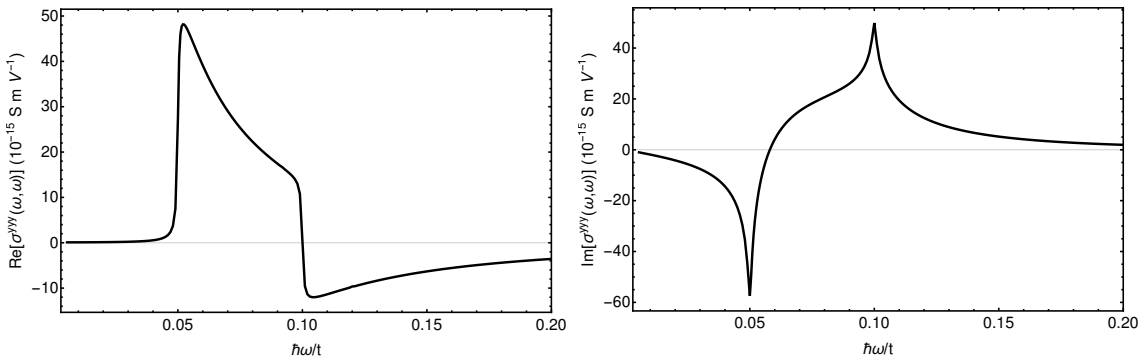
$$-2\sqrt{3}\,\cos\left(\frac{k_x\,a}{2}\right)\,\sin\left(\frac{\sqrt{3}\,k_y\,a}{2}\right) = 0 \tag{4.92}$$

which has several solutions:

$$k_x = 0 \wedge k_y = 0 \implies \epsilon_{\mathrm{vH}} = \sqrt{(9\,t)^2 + \left(\frac{\Delta}{2}\right)^2} \tag{4.93}$$

the chosen origin of the FBZ, corresponding to the conduction band energy maximum in Fig. 3.2,

$$k_x = \pm\frac{4\,\pi}{3\,a} \wedge k_y = 0 \implies \epsilon_{\mathrm{vH}} = \frac{\Delta}{2} \tag{4.94}$$



(a) (b)

Figure 4.5: Second order optical conductivity of gapped graphene with $\Delta = 0.1\,t$ and $\gamma = 5 \times 10^{-4}\,t$, at $T = 0\,\mathrm{K}$. The Fermi level resides in the band gap. The response displays sudden jumps at $\hbar\omega = \Delta/2$ and $\hbar\omega = \Delta$ in the real part (a) with accompanying changes in the imaginary part (b). Since the system is an insulator, there is no Drude peak at zero frequency and both real and imaginary parts vanish for $\omega \to 0^+$.

the notable Dirac points at $\mathbf{K}$ and $\mathbf{K}'$, vertices of the hexagonal FBZ, for which the energy is minimum, and

$$k_x = 0 \wedge k_y = \pm\frac{2\,\pi}{\sqrt{3}\,a} \implies \epsilon_{\mathrm{vH}} = \sqrt{t^2 + \left(\frac{\Delta}{2}\right)^2} \tag{4.95}$$

the middle points along the edges of the hexagonal FBZ, representing local maximums of the conduction band energy. By symmetry, it is possible to infer other states, with the same energy, that solve Eqs. 4.91 and 4.92. The complete set of Bloch states with van Hove singularities is depicted in Fig. 4.4b.

The last case (Eq. 4.95) is the one visible in Fig. 4.3, where the feature appears for $\hbar\omega = 2\,\epsilon_{\mathrm{vH}}(\Delta = 0) = 2\,t$. Curiously, the step now appears in the imaginary part and the divergence in the real part, reversing the roles.

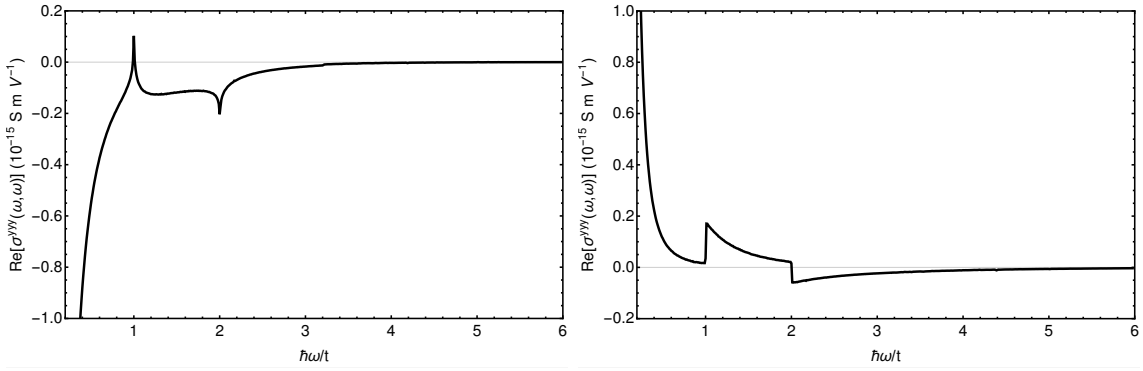
The analysis that occupies us here stays within the confines of the independent electron approximation and has its limitations in describing the observed optical and electronic properties of crystals. The high frequency van Hove singularity exemplifies this perfectly. It is known that, when accounting for the existence of excitons, whose treatment lies outside the scope of this thesis, the position of the van Hove singularity shifts considerably in the energy spectrum [90].

4.6.3 Second order conductivity

The presence of inversion symmetry in monolayer graphene dictates that all even orders of the optical response be identically zero. The absence of a second order response in particular was confirmed numerically with the minimal coupling method.

Inversion symmetry is broken by making the A- and B-site atoms nonequivalent, as in gapped graphene (Section 3.4). Gapped graphene has a nonzero second order conductivity, presented in Figs. 4.5 and 4.6.

In Fig. 4.5, the second order conductivity is governed by the electronic properties near the Dirac point. The real part displays a step when $2\,\hbar\omega = \Delta$ and another,



(a) (b)

Figure 4.6: Second order optical conductivity of gapped graphene with $\Delta = 0.1\,t$ and $\gamma = 5 \times 10^{-4}\,t$, at $T = 0\,\mathrm{K}$, for optical frequencies that probe regions of the FBZ beyond the Dirac point approximation. The nonlinear optical response has features at $\hbar\omega = 2\,\epsilon_{\mathrm{vH}}/2 \simeq t$ and $\hbar\omega = 2\,\epsilon_{\mathrm{vH}} \simeq 2\,t$.

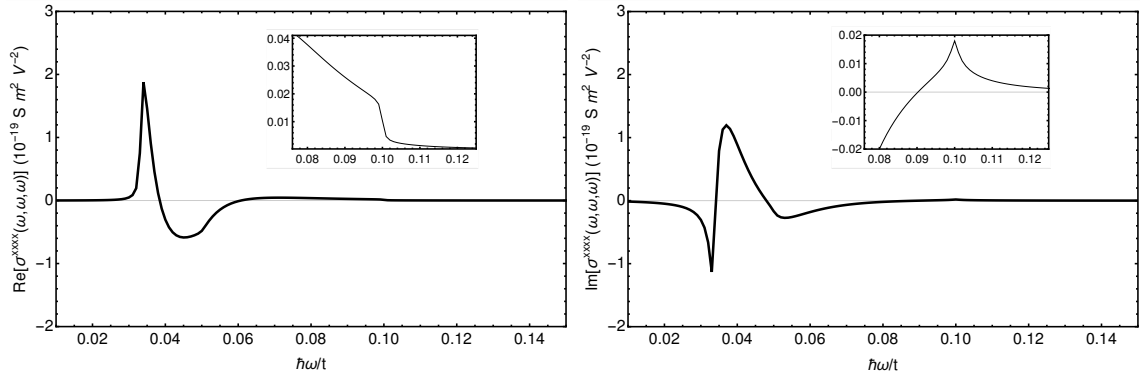
in the opposite direction, when $\hbar\omega = \Delta$, while the imaginary part has a negative logarithmic divergence at $2\,\hbar\omega = \Delta$ and a positive one at $\hbar\omega = \Delta$. Below the frequency whose double matches the gap, the real part of the conductivity again vanishes identically.

The ability of the minimal coupling algorithm to provide the optical response in an wide frequency range extends to the nonlinear response. As an example, the second order conductivity is depicted in Fig. 4.6 for higher frequencies. The behavior is similar to Fig. 4.3, but there are now two features, at $2\,\hbar\omega = 2\,\epsilon_{\mathrm{vH}}$ and at $\hbar\omega = 2\,\epsilon_{\mathrm{vH}}$, when the optical frequency or its double hit the van Hove singularity. This time, the features observed in the high frequency range are much smaller in magnitude than their low energy counterparts. A more detailed analysis on the nonlinear optical response of graphene beyond the Dirac point approximation can be found in [84].

The results of this subsection already demonstrate the applicability of the minimal coupling method to studies in nonlinear optics. As a curiosity, know that if the "conventional" velocity gauge treatment, described in Section 2.4.2, was (incorrectly) applied to this two-band model, the entire second order response would have returned zero. Furthermore, the first study of the third order optical conductivity of monolayer graphene over an extended frequency range [84] was done with, and made possible by, the minimal coupling method introduced here.

4.6.4 Third order conductivity

The third order conductivity of gapped graphene is in Fig. 4.7. At this point, a pattern starts to become noticeable. The real part of the conductivity contains discrete steps at the frequencies $3\,\hbar\omega = \Delta$, $2\,\hbar\omega = \Delta$ and $\hbar\omega = \Delta$, with alternating signs. Accompanying the steps in the real part, are the divergences in the imaginary part, which not only appear at the same frequencies, but always run in the opposite sense of the step-like increase/decrease, as in previous orders. Interestingly, the feature at $\hbar\omega = \Delta = 0.1\,t$, represented in the inset, is much smaller than the rest. Below $\hbar\omega = \Delta/3$, the real part vanishes, while the imaginary decays to zero in the



(a) (b)

Figure 4.7: Third order optical conductivity of gapped graphene with $\Delta = 0.1\,t$ and $\gamma = 5 \times 10^{-4}\,t$, at $T = 0\,\mathrm{K}$. The Fermi level resides in the band gap. The response displays sudden jumps at $\hbar\omega = \Delta/3$, $\hbar\omega = \Delta/2$ and $\hbar\omega = \Delta$ in the real part (a) with accompanying changes in the imaginary part (b). The inset zooms in at $\hbar\omega = \Delta$ to display the one-photon features, small in magnitude. Since the system is an insulator, there is no Drude peak at zero frequency and both real and imaginary parts vanish for $\omega \to 0^+$.

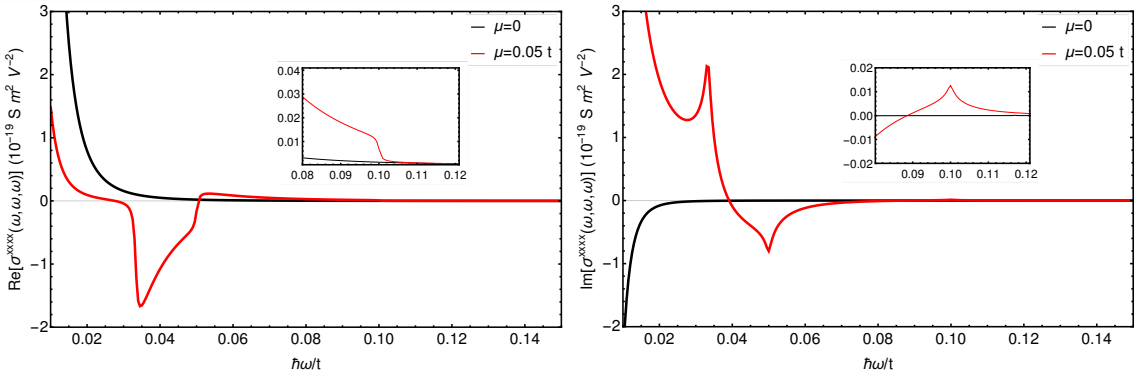
DC limit.

Closing the gap, all features vanish and the modulus of the conductivity decreases monotonically for increasing frequency, as seen in Fig. 4.8. When $\omega \to 0^+$, the nonlinear conductivity diverges, even though the concentration of carriers is zero, which raises the concern: if the third order DC conductivity of graphene is seemingly infinite, can we truly trust the perturbative results in this regime?

For undoped graphene, we cannot. The validity of perturbation theory relies on the interband matrix elements of the perturbation being smaller than $|\hbar\omega - \Delta|$, but in the absence of a band gap and when the optical frequency itself approaches zero, the plausibility of a perturbative treatment is lost and any results are to be mistrusted. If in linear order, the surprising answer $\sigma^{xx}(0) = \sigma_0$ could perhaps be accepted, $\sigma^{xxxx}(0) = \infty$ proves conclusively that the derived formulas are mathematically unsound in the DC limit. To put it simply, the series expansion of the current for a DC field does not converge. This point is easily missed in linear transport theory and is here made transparent by a nonlinear optics study.

These worries are removed by setting an effective gap, with $\mu \neq 0$. The third order optical response of doped graphene is traced in red in Fig. 4.8. Similarly to the gapped case, a sequence of steps and divergences are found when the optical frequency, its double or its triple match the effective gap. The strongest of these features is at $\hbar\omega = 2|\mu|/3$. Just as in the gapped case, the feature observed when the optical frequency matches the gap is, comparably, fairly weak.

The dispersion represented in Fig. 4.8 is in perfect agreement with the literature [56, 76], although there is a technical point to consider. When comparing the results of this thesis to works by other authors, it should be noted that the phenomenology adopted here, introduced in Section 2.5, is not the most common. This leads to a discrepancy between different phenomenological approaches only in a region of the order γ around the resonant frequencies ($2|\mu|/3\hbar$, $|\mu|/\hbar$ and $2|\mu|/\hbar$, in this case). This becomes irrelevant in the relaxation-free limit, but, for any finite γ,



(a) (b)

Figure 4.8: Third order optical conductivity of graphene ($\Delta = 0$) with $\gamma = 5\times10^{-4}\,t$, at $T = 0\,\mathrm{K}$. In black, the curves represent the case with no carriers, $\mu = 0$, and in red, the doped system with $\mu = 0.05\,t$, setting an effective gap at $2|\mu| = 0.1\,t$. The response displays sudden jumps at $\hbar\omega = 2|\mu|/3$, $\hbar\omega = 2|\mu|/2$ and $\hbar\omega = 2|\mu|$ in the real part (a) with accompanying changes in the imaginary part (b). The inset zooms in around $\hbar\omega = 2|\mu|$ to display the one-photon features, small in magnitude. For the doped system, the expected Drude peak is observed at lower frequencies.

the behavior is noticeably distinct in this narrow window, with our use of complex frequencies providing, arguably, more physically looking curves. This question is discussed in more detail in [56].

Lastly, there is the Drude peak in Fig. 4.8. As in the linear response, when there is a Fermi surface and negligible relaxation, a physical divergence is to be expected in the DC limit, where charge carriers flow in the direction of the electric field without impediment[8], ever increasing, never reaching a steady-state.

Outside the DC conductivity, but well below the effective gap, there is a range of frequencies where the real part of the optical response, linear and nonlinear, is entirely determined by Fermi surface properties (Chapter 5). In this region of the spectrum, Boltzmann equation treatments provide an intuitive, direct means to derive the optical response of the system. The first derivations of the third order conductivity of graphene followed this route [72, 74], with results that are reproduced by our more general calculations.

Aside from the consistency between theoretical calculations, often based on different techniques, what is desirable is to find agreement between theoretical predictions and experimental observation. In this front, the study of the nonlinear optical response of two-dimensional materials is still in its infancy and most measurements on graphene, in particular, are in significant disagreement with each other. On top of this, it shouldn't be expected that our naive independent electron picture fully captures the nonlinear physics involved in these experiments. On the other hand, the careful work of [78] has demonstrated that the most important aspects of the nonlinear response of graphene are, in fact, properly described by the dispersion behavior presented here, for sufficiently weak optical fields.

Among the possible reasons for the apparent (orders of magnitude) discrepancy between theory and experiments (prior to the agreement found in the work of Jiang *et al* [78]), the use of an effective Hamiltonian, the Dirac Hamiltonian, in all the

[8] In this idealized situation of infinitesimally small relaxation.

theoretical computations made by then was pointed out as a possible source of error [76, 77]. Since the minimal coupling method takes into account, and in fact *requires*, the entire band structure of graphene, this is now known to not be the case. The results derived in this chapter with the minimal coupling method, running the integration over every piece of the hexagonal First Brillouin Zone, reproduce previous calculations that adopted low energy effective descriptions around the Dirac points [76]. These calculations, even though they do not possess the versatility of the minimal coupling algorithm, are simpler and often provide analytical answers, providing insight on numerical results, as shall be demonstrated in the next chapter.

Chapter 5

Resonance-based analysis

In this chapter, our attention turns to the length gauge method. The structure of the electronic nonlinear optical conductivity will be elucidated by means of a detailed study of the two-band model. The nonlinear conductivity is decomposed as a sum of contributions related with different regions of the First Brillouin Zone, defined by single or multiphoton resonances. All contributions are written in terms of the same integrals, which contain all information specific to the particular model under study. In this way, ready-to-use formulas are provided that reduce the often tedious calculations of the second and third order optical conductivity to the evaluation of a small set of similar integrals. The treatment presented in this chapter is a direct generalization of the work in [91].

5.1 Motivation

While the minimal coupling method proves itself excellent for numerically assessing the dispersion of the nonlinear conductivity, especially in the context of tight binding models, it lacks the ability to handle low energy effective Hamiltonians. This is relevant, since these are usually the ones for which analytical answers can be found. By this, it is meant that low energy effective theories are typically simpler and one is able to integrate over reciprocal space and arrive at explicit analytical expressions for the linear and nonlinear conductivity.

For this task, the length gauge method is to be favoured. Still, the complexity is considerable and it can be a daunting task to try and derive a third order conductivity. Here, this complexity is brought down and calculations are made as straightforward as using Fermi's golden rule, though multiple times in sequence.

Another, perhaps just as important, goal of this chapter is to gather insight on the mathematical structure of the nonlinear conductivity. In the numerical results for monolayer graphene displayed in Section 4.6, features were found when resonance conditions were met at the gap, but no appropriate explanation was given for such. Additionally, there are other questions that we do not yet have the tools to answer: What possible features can be displayed by the dispersion of a nonlinear conductivity? Why are steps found in the real part and divergences in the imaginary part? These questions, among others, will be answered over the course of this chapter, by revisiting and inspecting in detail the length gauge perturbation theory of Aversa and Sipe [49].

In the length gauge, both intraband and interband transitions must be taken

into account and are transparently expressed in the structure of the perturbation theory, presented in Section 2.4.1, which contains as particular cases the dynamics of atomic systems and the free carriers single-band motion, but is more general, and complex, than either [49]. As we have discussed (Sections 2.3 and 2.4), the complexity stems from expressing the perturbation in terms of a position operator, which takes the form of a covariant derivative in the Bloch representation [49, 54, 57]. The successive application of derivatives as we move to higher orders in perturbation theory leads to unwieldy expressions for the nonlinear conductivity and lengthy calculations even for the very simplest models, the only ones for which analytical calculations are even attempted. Despite this, the results sometimes show surprising simplicity and structure. As an example, we shall see (Section 5.5) that the third order conductivity of the system of massless Dirac fermions found in monolayer graphene has the form[1] [76],

$$\sigma^{xxxx}(\omega,\omega,\omega) = \frac{C_0}{(\hbar\omega)^4}\left(-17\,G\left(\frac{\hbar\omega}{2|\mu|}\right) + 64\,G\left(\frac{2\hbar\omega}{2|\mu|}\right) - 45\,G\left(\frac{3\hbar\omega}{2|\mu|}\right)\right) \tag{5.1}$$

with

$$G(x) \equiv \Theta(|x|-1) + \frac{i}{\pi}\log\left|\frac{1-x}{1+x}\right| \tag{5.2}$$

where C_0 is a constant: $C_0 = \hbar\, v_F^2\, e^4/192$.

In physics, when elaborate and extensive calculations are required to derive simple and elegant results, it is sometimes a sign that a simpler and more insightful way to express the theory exists. This is the perspective we take here. Eq. 5.1 has the third order conductivity broken up into pieces that are relevant in different regions of the spectrum, depending on whether the energies of one, two or three photons are closer to matching the "effective gap" given by $2|\mu|$. This suggests that a resonance-based decomposition of the conductivity might be possible in general, leading to a more direct derivation of analytical results and easier interpretation of the underlying physics.

In exploring this possibility, we confine ourselves to the study of the two-band crystal, the solid-state analogue of the two-level atom. In similar spirit to the theoretical investigations of the nonlinear optics of two-level atoms during the seventies [3, 92], we expect a study of the two-band crystal to provide a firm foundation for later investigations of more general systems, to allow the central concepts to emerge more simply and to have a wide range of applicability, encompassing any situation where the incident photon frequencies connect a single pair of conduction (c) and valence (v) bands.

In Section 5.2, the central results from this study are presented. A decomposition of the nonlinear conductivity is made based on the possible resonances and ready-to-use formulas are provided that, for any two-band model, reduce the calculation of the linear, second and third order conductivity to the evaluation of one[2], two and six integrals over the FBZ, respectively. For systems that possess a Fermi surface (e.g. metals), there are two additional integrals to compute at each order. The integrals can sometimes be evaluated analytically in the relaxation-free limit, as

[1] In the relaxation-free limit $\gamma \to 0^+$.

[2] Two if the system is not topologically trivial.

described in Section 5.3. A brief discussion on how to obtain finite temperature results from a zero temperature calculation is included in Section 5.4, closing the exposition of the formalism. The ideas and tools developed in this chapter are illustrated with calculations for the system of massive and massless Dirac fermions present in monolayer graphene with and without a gap, respectively, in Section 5.5.

5.2 Photon resonances and the Fermi surface

In Section 2.4.1, it was shown that,

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) = \frac{i\,e^{n+1}}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\big[D^\beta,H_0\big]_{\mathbf{k}ba}}{\hbar\bar{\omega}_1+\cdots+\hbar\bar{\omega}_n-\Delta\epsilon_{\mathbf{k}ab}}\left[D^{\alpha_n},\ldots\frac{1}{\hbar\bar{\omega}_1-\Delta\epsilon}\circ\big[D^{\alpha_1},\rho_0\big]\ldots\right]_{\mathbf{k}ab} \tag{5.3}$$

From this concise expression, a more explicit form of the nonlinear conductivity can be derived by expanding out all the commutators and performing all the required differentiation (that follows from Eq. 2.52), resulting in a lengthy and rather cumbersome formula. This is the usual starting point in the literature, when computing the nonlinear optical response functions of semiconductors and other materials. Numerical integration is necessary, except for some cases where a low-energy description exists with very simple dispersion relation and eigenstates. For these systems, analytical calculations are sometimes possible, but still often rather long and complicated [75]. Here, we attempt to bring some simplicity and clarity to the structure of the nonlinear conductivity, by separating out terms whose resonances are located in different regions of the FBZ.

Since we are restricting ourselves to the analysis of a two-band system, there is a single (nonzero) energy difference in the denominators of Eq. 5.3, $\Delta\epsilon_{\mathbf{k}ab}=\pm\Delta\epsilon_{\mathbf{k}cv}$, allowing for a partial fraction decomposition into terms with a single denominator to be integrated, $(\hbar\bar{\omega}_1+\cdots+\hbar\bar{\omega}_i\pm\Delta\epsilon_{\mathbf{k}cv})^{-1}$ with $i\in\{1,\ldots,n\}$. These terms we denote by $\sigma_i^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_i,\ldots,\bar{\omega}_n)$ as they are associated with resonances involving an i number of photons[3]. We shall see later how the real part[4] of these contributions is entirely described by the properties of the crystal in the vicinity of regions of the FBZ where the resonance condition $\hbar\omega_1+\cdots+\hbar\omega_i-\Delta\epsilon_{\mathbf{k}cv}=0$ is met. Some terms will involve poles of higher orders, but these can reduced back to simple poles with an integration by parts or, equivalently, by making use of the identities in Appendix G. Finally, there will be terms where the application of the position operator resulted in derivatives of the Fermi-Dirac distribution. These terms will be treated separately and are denoted by $\sigma_F^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)$. An explicit application of the procedure outlined here can be found in Appendix F, where the second order conductivity is treated in detail. The resonance-based decomposition, stated generally, gives

[3] In due rigor, we work with a classical electromagnetic field and there are no photons present. It should, however, be clear that when a proper quantum treatment is made $\hbar\omega_i$ is the energy of an incident photon, thereby justifying the used nomenclature.

[4] In a time-reversal symmetric system.

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) = \sigma_F^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) + \sigma_1^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) + \cdots + \sigma_n^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) \tag{5.4}$$

The various pieces of Eq. 5.4 will be made explicit in the following sections, but it is useful to first inspect their structure in general terms. The one-photon contribution can always be written as

$$\sigma_1^{\alpha}(\bar{\omega}_1,\cdots,\bar{\omega}_n) = \sum_j \sum_p C_{1j}^{p}(\bar{\omega}_1,\cdots,\bar{\omega}_n)\,\Pi_j^{p(\alpha)}(\bar{\omega}_1) \tag{5.5}$$

where all tensor indices where condensed into one $\alpha \equiv \beta\alpha_1\ldots\alpha_n$ and p stands for permutation. The sum in p implies $p(\alpha)$ runs over all permutations of α, with a specific coefficient for each permutation applied. The coefficients $C_{1j}^{p}(\bar{\omega}_1,\cdots,\bar{\omega}_n)$ are specified in the following sections for the linear, second order and third order conductivities ($n = 1$, 2 and 3, respectively), where it is observed that most of these coefficients are zero, making only a small number of permutations necessary in practice. The coefficients are independent of the details of the system under consideration (they depend solely on the optical frequencies). All dependence on material properties in the sum of Eq. 5.5 is in the integrals Π_j^{α} that take the general form,

$$\Pi_j^{\alpha}(\bar{\omega}) = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{g_j^{\alpha}(\mathcal{A},\epsilon)_{\mathbf{k}ab}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.6}$$

with g_j^{α} as a set of functions, labeled by $j = 1, 2, \ldots$, that depend on the energies and their derivatives and on the non-abelian Berry connection $\mathcal{A}$ and its derivatives.

Similarly, for the two-photon contributions,

$$\sigma_2^{\alpha}(\bar{\omega}_1,\cdots,\bar{\omega}_n) = \sum_j \sum_p C_{2j}^{p}(\bar{\omega}_1,\cdots,\bar{\omega}_n)\,\Pi_j^{p(\alpha)}(\bar{\omega}_1+\bar{\omega}_2) \tag{5.7}$$

and the generalization is obvious at this point,

$$\sigma_i^{\alpha}(\bar{\omega}_1,\cdots,\bar{\omega}_n) = \sum_j \sum_p C_{ij}^{p}(\bar{\omega}_1,\cdots,\bar{\omega}_n)\,\Pi_j^{p(\alpha)}(\bar{\omega}_1+\cdots+\bar{\omega}_i) \tag{5.8}$$

Since all contributions involve a combination of the same integrals with a changing argument, the calculation of the nonlinear conductivity is reduced to the evaluation of the integrals in Eq. 5.6. The complexity and number of integrals to be evaluated increases with the order n of the nonlinear conductivity, but they always retain the general form of Eq. 5.6 for some g function. For finite γ, numerical integration will invariably be required. In Section 5.3, we analyze the limit of vanishing relaxation, where analytical results are accessible.

It is worth noting at this point that the conductivity in Eq. 5.4 is not symmetrized. It follows from the definition in Eq. 1.9 that only the portion of the nonlinear conductivity that respects intrinsic permutation symmetry is physical [3] (see the discussion at the end of Section 2.2). When permutations of Eq. 5.4 are properly accounted for, there will not only be an one-photon contribution associated to the resonance $\hbar\omega_1 = \Delta\epsilon_{\mathbf{k}cv}$, but also $\hbar\omega_2 = \Delta\epsilon_{\mathbf{k}cv}$ and so on. Likewise for contributions associated with higher numbers of photons. Having this in mind, the formulas presented for the nonlinear conductivity in this chapter will nonetheless be

left, for the most part, unsymmmetrized. Symmetrizing the conductivity is a trivial, if cumbersome, procedure and adds little to the discussion here.

The terms described so far give a complete description for any two-band model that is used to describe an insulator or a cold semiconductor. But for systems with free charge carriers prior to optical excitation, there is an additional contribution[5],

$$\sigma_F^{\alpha}(\bar{\omega}_1, \cdots, \bar{\omega}_n) = \sum_{X=A,B,\ldots} \sum_{p} C_X^{p}(\bar{\omega}_1, \cdots, \bar{\omega}_n)\, F_X^{p(\alpha)} \tag{5.9}$$

where the integrals have a different structure than before,

$$F_X^{\beta\alpha_1\cdots\alpha_n} = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a g_X^{\alpha_1\ldots\alpha_n}(\mathcal{A}, \epsilon)_{\mathbf{k}aa}\; \partial^{\beta} f_{\mathbf{k}a} \tag{5.10}$$

It is evident from the presence of derivatives of Fermi functions in Eq. 5.10 that these integrals are determined by the properties of the Fermi surface. More surprising is the absence of any frequency dependence. All dispersion in σ_F comes from the C coefficients, which are the same for every two-band system. This leads us to an important result: *the dispersion of the contributions from the Fermi surface is given by an universal family of functions of frequency* (Eq. 5.9), obtained through linear combinations of the C_X's. The particular linear combination observed is set by the integrals F_X. Being dictated by Fermi surface properties, they are particularly dependent on carrier concentration and can therefore be tuned by doping.

In the following sections, Eqs. 5.5-5.10 are made explicit for the linear, second and third order conductivities. Ready-to-use formulas are presented that reduce the calculation of the nonlinear conductivities to the evaluation of a minimal number of integrals over the FBZ (Eqs. 5.19-5.21, 5.26-5.29, 5.35-5.42).

5.2.1 Linear order

The derivation

The linear conductivity is simple enough that the suggested resonance-based decomposition can be treated in detail and its derivation included here. The analogous treatment for the second order conductivity takes greater effort and is relegated to Appendix F. The derivation of the third order conductivity decomposition is entirely omitted, as it is too extensive, and only the final expressions are presented.

Expanding the commutators (Eqs. 2.53 and 2.55) in the linear conductivity, as they appear in the length gauge version (Eq. 2.61),

$$\begin{aligned}\sigma^{\beta\alpha_1}(\bar{\omega}_1) &= -\frac{i\,e^2}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\left[D^{\beta}, H_0\right]_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} \left[D^{\alpha_1}, \rho_0\right]_{\mathbf{k}ab} \\ &= -\frac{i\,e^2}{\hbar^2\,\bar{\omega}_1} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \left(\partial^{\beta}\epsilon_{\mathbf{k}a}\right)\left(\partial^{\alpha_1} f_{\mathbf{k}a}\right) + \frac{i\,e^2}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\Delta\epsilon_{\mathbf{k}ab}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba}\end{aligned} \tag{5.11}$$

[5] The validity of Eq. 5.9 has been checked by the author only up to third order.

$$= -\frac{i\,e^2}{\hbar^2\,\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a\left(\partial^\beta\epsilon_{\mathbf{k}a}\right)\left(\partial^{\alpha_1}f_{\mathbf{k}a}\right)+i\,e^2\,\bar{\omega}_1\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathcal{A}^\beta_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}\Delta f_{\mathbf{k}ba}$$
$$-\frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^\beta_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba} \tag{5.12}$$

where the last term can be rewritten with the use of Eq. C.6 from Appendix C,

$$-\frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^\beta_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba}=-\frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}[\mathcal{A}^\beta,\mathcal{A}^{\alpha_1}]_{\mathbf{k}aa}\,f_{\mathbf{k}a}$$
$$=-\frac{e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a\mathcal{F}^{\beta\alpha_1}_a\,f_{\mathbf{k}a} \tag{5.13}$$

Replacing back[6] in Eq. 5.12,

$$\sigma^{\beta\alpha_1}(\bar{\omega}_1)=-\frac{i\,e^2}{\hbar^2\,\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a\left(\partial^\beta\epsilon_{\mathbf{k}a}\right)\left(\partial^{\alpha_1}f_{\mathbf{k}a}\right)-\frac{e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a\mathcal{F}^{\beta\alpha_1}_a\,f_{\mathbf{k}a}$$
$$+i\,e^2\,\bar{\omega}_1\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathcal{A}^\beta_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}\Delta f_{\mathbf{k}ba} \tag{5.14}$$

The first term in Eq. 5.14 contains the derivative of a Fermi function and can be recognized as a Fermi surface integral of the type in Eq. 5.10. The second term fits the mold of Eq. 5.6, identifiable by the one-photon resonance at $\hbar\omega_1=\Delta\epsilon$. Strangely, the last term falls in neither category. It is somewhat unfortunate, but we have already encountered a deviation from the pattern laid out at the beginning of this section.

The last term in Eq. 5.14 is a global term, that cannot be fully captured by the properties of any given subregion of the FBZ. This type of term exists only at linear order and forms the single exception to the general scheme proposed in Eq. 5.4. It is known as the anomalous Hall conductivity [93] and is responsible for the intrinsic quantum Hall effect: it provides a non-vanishing current in the absence of a magnetic field and a Fermi surface.

In two-dimensional crystals, if the Fermi level lies in the gap and $T=0\,\mathrm{K}$, the sum in Eq. 5.13 runs over the occupied bands only and returns a topological invariant: the sum of the Berry curvature flow associated to the occupied bands. This quantity is quantized, since

$$\int\frac{d^2\mathbf{k}}{(2\pi)^2}\sum_v\mathcal{F}^{xy}_v=\frac{N}{2\,\pi} \tag{5.15}$$

where the band index v considers only the occupied valence bands and N is the Chern number, an integer.

If the Fermi level crosses a band, there is an additional non-integer part that can, by performing an integration by parts, be identified as a Fermi surface property. For this reason, and frankly for notational convenience, the term in Eq. 5.13

[6] Had the starting point been Eq. D.6 instead of Eq. 2.61, the same result would have been produced, without the use of Eq. C.6.

will be classified under the label "Fermi surface integral", which is appropriate for topologically trivial systems ($N = 0$).

For the most part, the anomalous contribution will not be crucial to the ideas of this chapter. It vanishes in the presence of time-reversal symmetry (Section 6.6) and does not affect the nonlinear optical properties of the crystal, our main interest.

The decomposition

In linear order, the resonance-based decomposition of Eq. 5.4 falls into the familiar intra- and interband separation

$$\sigma^{\beta\alpha_1}(\bar{\omega}_1) = \sigma_F^{\beta\alpha_1}(\bar{\omega}_1) + \sigma_1^{\beta\alpha_1}(\bar{\omega}_1) \tag{5.16}$$

with

$$\sigma_F^{\beta\alpha_1}(\bar{\omega}_1) = \frac{i\,e^2}{\hbar}\left(-\frac{1}{\hbar\bar{\omega}_1}\,F_A^{\beta\alpha_1} + i\,F_B^{\beta\alpha_1}\right) \tag{5.17}$$

$$\sigma_1^{\beta\alpha_1}(\bar{\omega}_1) = \frac{i\,e^2}{\hbar}\,\hbar\bar{\omega}_1\,\Pi_1^{\beta\alpha_1}(\bar{\omega}_1) \tag{5.18}$$

The integrals required to obtain the linear response are well-known and particularly simple. There is one integral associated with an interband resonance and two related to the Fermi surface (including the anomalous contribution from before),

$$F_A^{\beta\alpha_1} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a (\partial^{\alpha_1}\epsilon_{\mathbf{k}a})\,(\partial^{\beta} f_{\mathbf{k}a}) \tag{5.19}$$

$$F_B^{\beta\alpha_1} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \mathcal{F}_a^{\beta\alpha_1}\, f_{\mathbf{k}a} \tag{5.20}$$

$$\Pi_1^{\beta\alpha_1}(\bar{\omega}_1) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_1}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.21}$$

5.2.2 Second order

In second order, Eq. 5.4 gives

$$\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_F^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) + \sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) + \sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) \tag{5.22}$$

Each contribution will be inspected separately. Their derivation is presented in detail in Appendix F. In what follows, the C coefficients in Eq. 5.8 and the g functions in Eqs. 5.6 and 5.10 for $n = 2$ will be made explicit.

Starting with the Fermi surface contribution,

$$\sigma_F^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \frac{i\,e^3}{\hbar}\left(\frac{i}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2}\,F_A^{\beta\alpha_1\alpha_2} + \frac{1}{\hbar\bar{\omega}_1}\,F_B^{\alpha_1\alpha_2\beta}\right) \tag{5.23}$$

The one-photon contribution,

$$\frac{\hbar}{i\,e^3}\,\sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}\,\Pi_1^{\alpha_2\alpha_1\beta}(\bar{\omega}_1) + \frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1) + \frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\,\Pi_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1) \tag{5.24}$$

Finally, the two-photon contribution,

$$\frac{\hbar}{i\,e^3}\,\sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = -\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) + \frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) \tag{5.25}$$

In second order, there are two integrals from interband resonances and two from the Fermi surface,

$$F_A^{\beta\alpha_1\alpha_2} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \left(\partial^{\alpha_1}\partial^{\alpha_2}\epsilon_{\mathbf{k}a}\right)\left(\partial^{\beta} f_{\mathbf{k}a}\right) \tag{5.26}$$

$$F_B^{\beta\alpha_1\alpha_2} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \mathcal{F}_a^{\alpha_1\alpha_2}\left(\partial^{\beta} f_{\mathbf{k}a}\right) \tag{5.27}$$

$$i\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_1}\left(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab}\right)}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.28}$$

$$i\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab;\alpha_2}^{\alpha_1}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.29}$$

where we made use of the generalized derivative notation introduced by Aversa and Sipe [49]: $\mathcal{O}_{\mathbf{k}ab;\alpha} \equiv \left(\delta_{ab}\,\partial^{\alpha} - i(\mathcal{A}_{\mathbf{k}aa}^{\alpha} - \mathcal{A}_{\mathbf{k}bb}^{\alpha})\right)\mathcal{O}_{\mathbf{k}ab}$.

We chose, in Eq. 5.20, to label the anomalous contribution to the linear conductivity as F_B, and place it under the class of Fermi surface contributions. After inspection of Eqs. 5.26 and 5.27, it is evident why. They form a natural extension of their linear counterparts in Eqs. 5.19 and 5.20, with one more derivative applied to the Fermi function in F_B or to the band dispersion in F_A. This will continue to be the case for higher orders in perturbation theory.

F_A is the purely intraband contribution. That is, F_A is to be identified as the contribution that would be present even in a system where all interband transitions are neglected. In a one-band model, it is the only term that survives.

While the purely intraband contributions can be derived to any order without the need for a perturbation theory as complex as the one discussed here, the inclusion of interband transitions brings new Fermi surface integrals, involving the Berry curvature and its derivatives. For time-reversal symmetric systems, F_A vanishes (Section 6.6) and it is possible to isolate F_B. In this case, the Fermi surface contribution at second order, $\sigma_F^{\beta\alpha_1\alpha_2}$, is a probe of the Berry curvature around the Fermi surface.

Another curious result that follows from Eq. 5.27 is that we can expect the Fermi surface contribution to be absent when the current is measured in the direction of the optical fields: $\sigma_F^{\beta\alpha_1\alpha_2} = 0$ when $\beta = \alpha_1 = \alpha_2$. As a consequence, *no Drude peaks should be found in the diagonal tensor elements of the second order conductivity, for time-reversal symmetric systems.* This is true even in the presence of free carriers, found in metals or doped semiconductors.

Finally, we note that the contributions that probe the FBZ beyond the Fermi surface, the one- and two-photon contributions, are determined by the integrals in Eqs. 5.28 and 5.29 which are known in the literature of the nonlinear optics of solids for their relation to the injection and shift currents in semiconductors [5], respectively.

5.2.3 Third order

A considerable jump in complexity occurs when we move to third order. Once again, we start with Eq. 5.4,

$$\begin{aligned}\sigma^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \sigma_F^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) + \sigma_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) \\ + \sigma_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) + \sigma_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3)\end{aligned} \tag{5.30}$$

and write out all contributions explicitly.

The Fermi surface contribution,

$$\sigma_F^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \frac{i\,e^4}{\hbar}\left(\frac{1}{6\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3}\,F_A^{\beta\alpha_1\alpha_2\alpha_3} + \frac{i}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}\,F_B^{\alpha_2\beta\alpha_1\alpha_3}\right) \tag{5.31}$$

The one-photon contribution,

$$\begin{aligned}&\frac{\hbar}{i\,e^4}\,\sigma_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \\ &\frac{\hbar\bar{\omega}_{23}}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{123}}\,\Pi_1^{\alpha_1\alpha_2\alpha_3\beta}(-\bar{\omega}_1) + \frac{\hbar\bar{\omega}_{13}}{\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\alpha_1\alpha_2\beta\alpha_3}(-\bar{\omega}_1) - \frac{1}{\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_3}\,\Pi_2^{\alpha_2\alpha_1\alpha_3\beta}(\bar{\omega}_1) \\ &- \frac{\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1) - \frac{1}{(\hbar\bar{\omega}_3)^2\,\hbar\bar{\omega}_{123}}\,\Pi_3^{\alpha_2\alpha_1\alpha_3\beta}(\bar{\omega}_1) - \frac{\hbar\bar{\omega}_{123}}{2\,(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1) \\ &- \frac{\hbar\bar{\omega}_1}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3}\Pi_4^{\alpha_1\alpha_2\alpha_3\beta}(-\bar{\omega}_1) + \frac{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_{13}\,\hbar\bar{\omega}_{23}}\,\Pi_5^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_1) - \frac{(\hbar\bar{\omega}_1)^2\,\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_{13}}\,\Pi_5^{\alpha_2\alpha_3\alpha_1\beta}(\bar{\omega}_1)\end{aligned} \tag{5.32}$$

with the abbreviations $\bar{\omega}_{ij} \equiv \bar{\omega}_i + \bar{\omega}_j$ and $\bar{\omega}_{123} \equiv \bar{\omega}_1 + \bar{\omega}_2 + \bar{\omega}_3$.

The two-photon contribution,

$$\begin{aligned}&\frac{\hbar}{i\,e^4}\,\sigma_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \\ &- \frac{\hbar\bar{\omega}_{12}}{(\hbar\bar{\omega}_2)^2\,\hbar\bar{\omega}_3}\,\Pi_1^{\alpha_1\alpha_3\beta\alpha_2}(-\bar{\omega}_{12}) - \frac{\hbar\bar{\omega}_{12}}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_{123}}\,\Pi_1^{\alpha_3\alpha_1\alpha_2\beta}(\bar{\omega}_{12}) - \frac{\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{12}) \\ &+ \frac{1}{(\hbar\bar{\omega}_1)^2\,\hbar\bar{\omega}_{123}}\Pi_3^{\alpha_3\alpha_2\alpha_1\beta}(\bar{\omega}_{12}) + \frac{\hbar\bar{\omega}_{123}}{(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{12}) - \frac{(\hbar\bar{\omega}_{12})^2}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3}\,\Pi_6^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_{12})\end{aligned} \tag{5.33}$$

Finally, the three-photon contribution,

$$\frac{\hbar}{i\,e^4}\,\sigma_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \\ + \frac{\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) + \frac{\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) \\ - \frac{\hbar\bar{\omega}_{123}}{2\,(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_{13}\,\hbar\bar{\omega}_{23}}\,\Pi_5^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_{123}) \tag{5.34}$$

In third order, there are six integrals from interband resonances and two from the Fermi surface,

$$F_A^{\beta\alpha_1\alpha_2\alpha_3} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \left(\partial^{\alpha_1}\partial^{\alpha_2}\partial^{\alpha_3}\epsilon_{\mathbf{k}a}\right)\left(\partial^{\beta} f_{\mathbf{k}a}\right) \tag{5.35}$$

$$F_B^{\beta\alpha_1\alpha_2\alpha_3} \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \left(\partial^{\alpha_3}\mathcal{F}_a^{\alpha_1\alpha_2}\right)\left(\partial^{\beta} f_{\mathbf{k}a}\right) \tag{5.36}$$

$$\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab;\alpha_2}^{\alpha_1}\left(\partial^{\alpha_3}\Delta\epsilon_{\mathbf{k}ab}\right)}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.37}$$

$$\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_1}\left(\partial^{\alpha_2}\partial^{\alpha_3}\Delta\epsilon_{\mathbf{k}ab}\right)}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.38}$$

$$\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_1}\left(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab}\right)\left(\partial^{\alpha_3}\Delta\epsilon_{\mathbf{k}ab}\right)}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.39}$$

$$\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ab;\alpha_2\alpha_3}^{\alpha_1}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.40}$$

$$\Pi_5^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba}^{\beta}\,\mathcal{A}_{\mathbf{k}ba}^{\alpha_1}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_2}\,\mathcal{A}_{\mathbf{k}ab}^{\alpha_3}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.41}$$

$$\Pi_6^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}_{\mathbf{k}ba;\alpha_1}^{\beta}\,\mathcal{A}_{\mathbf{k}ab;\alpha_3}^{\alpha_2}}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{5.42}$$

There are, in the expressions of this section, two special cases worth addressing. These concern two types of systems that have been well studied in the past: metals and atomic/molecular gases.

First, the case of a metallic system at low frequencies. Consider that the Fermi level crosses a single band of our electronic system. In the limit of low frequencies, it is assumed that the Fermi surface contributions dominate and the other bands can be neglected. This produces the same conductivities that could be derived via a Boltzmann equation approach, but naturally includes the contributions involving the Berry curvature (such as the intrinsic Hall conductivity in linear order).

The well-known, much easier, perturbative results for atomic/molecular systems can be retrieved from the previous expressions for the linear and nonlinear conductivities by simply removing any dependence on the Bloch vector $\mathbf{k}$ (and discarding

the integration as well). The Fermi surface contributions automatically disappear and the non-abelian Berry connection, independent of $\mathbf{k}$, takes then the role of the dipole matrix elements (aside from a multiplicative constant). Most integrals vanish in this case. For instance, at third order, only Π_4, Π_5 and Π_6 survive. Notice that generalized derivatives do not return automatically zero, unless the diagonal elements of the position operator ($\mathcal{A}_{\mathbf{k}aa}$) are themselves zero or identical for both bands. If the position operator matrix elements have no diagonal matrix elements, a common approximation [3], then the Π_5 integrals are all that remains and describe the expected third order optical response of a gaseous system.

In crystals, all the listed integrals must be evaluated. The complexity is attenuated only by whatever symmetries the specific crystalline structure may possess. Crystal symmetry reduces the number of independent elements in the conductivity tensor, as was seen in Section 4.6.1 for graphene. Since this type of symmetry relates to the tensor indices and must be valid for any set of optical frequencies, it must also be obeyed by each Π_j^α individually. For graphene, for instance, all the integrals in Eqs. 5.35-5.42 will necessarily have the same symmetry properties presented in Eqs. 4.84-4.88 for the third order conductivity. This is of great practical relevance, as it limits the required number of integrals to compute to obtain the complete nonlinear conductivity.

There are other symmetries that are of a more general character and these will be addressed in Chapter 6. It is worth mentioning at this point, though, that gauge invariance is transparently expressed in the integrands of the Π_j^α integrals. Recall the discussion at the end of Section 2.3: the nonlinear conductivity must be invariant under a change of phases of the Bloch functions $\psi_{\mathbf{k}a} \to e^{i\phi_{\mathbf{k}a}}\psi_{\mathbf{k}a}$. It is as a direct consequence of this symmetry that all products of off-diagonal matrix elements of $\mathcal{A}$ appear in conjugate pairs (one $\mathcal{A}_{\mathbf{k}cv}$ and one $\mathcal{A}_{\mathbf{k}vc}$) and generalized instead of regular derivatives of $\mathcal{A}$ appear in Eqs. 5.35-5.42.

The type of resonance-based analysis presented in this section could be continued for successively higher order nonlinear conductivities. Unlike the minimal coupling method, the generalization is not immediate and requires finding new expressions at each order, with increasingly more elaborate derivations. The number of integrals continuously increases with n as more derivatives or generalized derivatives are applied in perturbation theory and more ways exist to arrange them in the numerators of the $\Pi_j^{\beta\alpha_1\cdots\alpha_n}$ integrals.

5.3 Relaxation-free limit

In Section 2.5, complex frequencies were introduced as a means to allow the physical system to relax. In this section, the limit of vanishing relaxation and real frequencies is considered. It is in this limit that the optical response is more physically transparent and analytical results more accessible.

However, there are some subtleties in implementing this. For certain regions of frequency space, the nonlinear conductivity will *diverge* in the relaxation-free limit. Specifically, the nonlinear conductivity may diverge in the relaxation-free limit when a subset of the optical frequencies adds to zero. This includes the case of DC fields ($\omega_i = 0$) or DC currents ($\omega_1 + \cdots + \omega_n = 0$).

Fundamentally, this is due to the existence of regions of the FBZ where more than one of the resonance conditions is obeyed simultaneously (two or more de-

nominators in the symmetrized version of Eq. 5.3 are resonant at those **k**-points). Geometrically, this can be visualized by representing in the FBZ the set of points where the conditions $\omega_1 - \Delta\epsilon_{\mathbf{k}cv} = 0$ or $\omega_1 + \cdots + \omega_i - \Delta\epsilon_{\mathbf{k}cv} = 0$ $(i = 2, ..., n)$ or any of their permutations are verified and checking for overlapping regions (for example, in Fig. 5.1b for there to be the possibility of a divergence two of the contours would have to cross or altogether overlap).

When using the equations of Section 5.2, this difficulty is realized in the relaxation-free limit of Eq. 5.8,

$$\begin{aligned}\lim_{\gamma\to 0^+} \sigma_i^{\alpha}(\bar{\omega}_1, \cdots, \bar{\omega}_n) &= \lim_{\gamma\to 0^+} \sum_{j,p} C_{ij}^{p}(\bar{\omega}_1, \cdots, \bar{\omega}_n)\, \Pi_j^{p(\alpha)}(\bar{\omega}_1 + \cdots + \bar{\omega}_i) \\ &= \sum_{j,p} C_{ij}^{p}(\omega_1, \cdots, \omega_n) \lim_{\gamma\to 0^+} \Pi_j^{p(\alpha)}(\bar{\omega}_1 + \cdots + \bar{\omega}_i) \end{aligned} \tag{5.43}$$

The validity of this passage hangs on the convergence of the C coefficients in the relaxation-free limit, which in turn relies on there existing no cancelling subset of frequencies in $\{\omega_1, \ldots, \omega_n\}$.

Henceforth, we will assume this to be the case. We emphasize that the validity of the equations in Section 5.2 always holds, even in these regions of frequency space. But the analysis of the relaxation-free limit in these situations, which include several nonlinear optical effects of interest (photovoltaic effects, DC field-induced second order response, electro-optic effects,...), albeit possible, requires greater care and lies outside the scope of this thesis. Indeed, the study of the divergences of the nonlinear optical conductivity is a topic of current research interest [6, 8, 94].

From Eq. 5.43, the evaluation of the relaxation-free limit of the nonlinear conductivity has been reduced to evaluating the relaxation-free limit of the integrals Π_j^{α}. This is a direct application of Eq. 2.78,

$$\lim_{\gamma\to 0^+} \Pi_j^{\alpha}(\bar{\omega}) = \pi \left(\mathcal{H}_j^{\alpha}(\omega) - i\,\mathcal{I}_j^{\alpha}(\omega)\right) \tag{5.44}$$

where

$$\mathcal{I}_j^{\alpha}(\omega) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} g_j^{\alpha}(\mathcal{A}, \epsilon)_{\mathbf{k}ab}\, \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega - \Delta\epsilon_{\mathbf{k}ab}\right) \tag{5.45}$$

$$\mathcal{H}_j^{\alpha}(\omega) \equiv \frac{1}{\pi} \fint \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{g_j^{\alpha}(\mathcal{A}, \epsilon)_{\mathbf{k}ab}}{\hbar\omega - \Delta\epsilon_{\mathbf{k}ab}}\, \Delta f_{\mathbf{k}ba} \tag{5.46}$$

Occasionally, the Π_j^{α} integrals occur with a minus sign in the argument (see Eqs. 5.24 and Eqs. 5.32-5.33). In this case, the pole in the Argand plane is approached from below, instead of above. Extending Eq. 5.44 to include this case

$$\lim_{\gamma\to 0^+} \Pi_j^{\alpha}(\pm\bar{\omega}) = \pi \left(\mathcal{H}_j^{\alpha}(\pm\omega) \mp i\,\mathcal{I}_j^{\alpha}(\pm\omega)\right) \tag{5.47}$$

As an example, let us consider the linear conductivity. Neglecting the Fermi surface contribution (Eq. 5.18),

$$\sigma^{\beta\alpha_1}(\omega_1) \equiv \lim_{\gamma\to 0^+} \sigma^{\beta\alpha_1}(\bar{\omega}_1) = \pi\, i\, e^2\, \omega_1 \left(\mathcal{H}_1^{\beta\alpha_1}(\omega_1) - i\,\mathcal{I}_1^{\beta\alpha_1}(\omega_1)\right) \tag{5.48}$$

where

$$\mathcal{I}_1^{\beta\alpha_1}(\omega_1) = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\, \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right) \tag{5.49}$$

$$\mathcal{H}_1^{\beta\alpha_1}(\omega_1) = \frac{1}{\pi} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab}}{\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}}\, \Delta f_{\mathbf{k}ba} \tag{5.50}$$

The integral $\mathcal{H}_1$ can be computed through Eq. 5.50 or from applying a Hilbert transform to $\mathcal{I}_1$. This is true in general, as it is easily proven from Eqs. 5.45 and 5.46 that

$$\mathcal{H}_j^{\alpha}(\omega) = -\frac{1}{\pi} \fint_{-\infty}^{+\infty} \frac{\mathcal{I}_j^{\alpha}(\omega')}{\omega' - \omega}\, d\omega' \tag{5.51}$$

The entire linear response of an insulator or a cold semiconductor is obtained, for negligible relaxation, from a calculation of the function in Eq. 5.49 and its Hilbert transform. This is by no means new information in linear optics: the integral in Eq. 5.49 is nothing more than Fermi's golden rule and the Hilbert transform from $\mathcal{I}_1$ to $\mathcal{H}_1$ equates to the usual derivation of refraction from absorption by the Kramers-Krönig relations. Here, however, this method has been generalized to any order.

In Section 6.6, it is shown that the integrals $\mathcal{I}(\omega)$ and $\mathcal{H}(\omega)$ are always real for systems that possess time-reversal symmetry. In this case, the real part of the nonlinear conductivity is given by the $\mathcal{I}$ integrals and the imaginary part by the $\mathcal{H}$ integrals[7]. This can be seen by noting that the coefficients C that multiply the integrals in Eq. 5.43 are purely imaginary in the relaxation-free limit.

The association of the real part of the nonlinear conductivity with the $\mathcal{I}$ integrals is to be expected if one recalls that the real part is usually the one responsible for optical absorption. If all relevant optical frequencies lie below the band gap, there must be no absorption and therefore the real part of the nonlinear conductivity of a cold semiconductor is required to vanish in this frequency range. This is guaranteed to be the case with the integrals in Eq. 5.45.

We conclude from the analysis of this section that, if we wish to evaluate the real part of a nonlinear conductivity, we have only to evaluate simple integrals with the form of Eq. 5.45 for the appropriate g functions, listed in Eqs. 5.19-5.21, 5.26-5.29, 5.35-5.42. These contributions are well localized in the FBZ and run over regions that satisfy the resonance conditions.

The imaginary part of a nonlinear conductivity follows directly by performing Hilbert transforms of the integrals computed for the real part. We note that we are not performing a Hilbert transform of the entire nonlinear conductivity as it is traditionally done in the nonlinear Kramers-Krönig relations [95]. We are not even applying a Hilbert transform to the different contributions in Eq. 5.4 to move from their real to imaginary parts[8]. The passage from the real to the imaginary part of the nonlinear optical conductivity is more subtle and is made directly through the integrals in Eq. 5.44. In this way, the imaginary part, which unlike the real part

[7] Once again, excluding the regions of frequency space where subsets of frequencies add to zero.

[8] This would not work as the real and imaginary parts of each individual contribution are not simply related by a Hilbert transform.

is not described by any restricted region of the FBZ, is made more accessible than what would perhaps be expected.

Finally, we point out that for Fermi surface contributions there is no need for the relaxation-free limit, since the F_X integrals are frequency independent. To attempt an analytical calculation it is necessary only to set the temperature to zero, in which case

$$F_X^{\beta\alpha_1\cdots\alpha_n} = -\int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a g_X^{\alpha_1\ldots\alpha_n}(\mathcal{A},\epsilon)_{\mathbf{k}aa} \left(\partial^\beta \epsilon_{\mathbf{k}a}\right) \delta(\mu - \epsilon_{\mathbf{k}a}) \tag{5.52}$$

This integral has a similar structure to the integrals in Eq. 5.45, with the chemical potential taking the role of a frequency.

For the previous integrals (Eq. 5.45), it is equally helpful to consider the system to be at zero temperature to make calculations tractable. However, once the nonlinear conductivity is calculated for $T = 0$, it is possible to obtain the answers at finite temperature with a few tricks [77].

5.4 Finite temperature

In Section 4.6, the numerical computations, employing the minimal coupling method, were all performed for $T = 0$, even though the algorithm accommodated the possibility of computing room or even high temperature conductivities. The reason was that, at low temperatures, the features in the nonlinear conductivity are more pronounced and its structure more transparent. This is demonstrated by the results in this section, where the effect of temperature on the conductivity is discussed.

In the length gauge also, more often than not, analytical calculations are made for zero temperature, where they are considerably simpler. It is, however, possible to quickly arrive at finite temperature results from zero temperature calculations of the nonlinear conductivity by making use of a relation first introduced in [77].

The key idea is to recognize that the entire chemical potential dependence of the $\Pi_i^\alpha(\bar{\omega})$ and F_X^α integrals stems from the Fermi-Dirac distribution. Starting with the former type of integrals, we make all arguments of the distribution explicit,

$$f(\epsilon,\mu,T) \equiv \frac{1}{e^{(\epsilon-\mu)/k_BT}+1} \tag{5.53}$$

At zero temperature,

$$f(\epsilon,\mu,0) = \Theta(\mu-\epsilon) \tag{5.54}$$

The finite temperature and zero temperature distributions are related by

$$f(\epsilon,\mu,T) = \int_{-\infty}^{+\infty} f(\mu',\mu,T)\,\delta(\mu'-\epsilon)d\mu' = \int_{-\infty}^{+\infty} f(\mu',\mu,T)\,\partial_{\mu'} f(\epsilon,\mu',0)\,d\mu' \tag{5.55}$$

Since no other objects in $\Pi_i^\alpha(\bar{\omega})$ depend on the chemical potential, Eq. 5.55 translates directly to

$$\sigma_i^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n;\mu,T) = \int_{-\infty}^{+\infty} f(\mu',\mu,T)\ \ \partial_{\mu'}\sigma_i^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n;\mu',0)\,d\mu' \tag{5.56}$$

where we made the chemical potential and temperature dependencies explicit.

By a similar argument, σ_F is proven to also obey Eq. 5.56. Joining all contributions, we arrive at a general relation between the zero and finite temperature nonlinear conductivity,

$$\begin{aligned}\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n;\mu,T) &= \int_{-\infty}^{+\infty} f(\mu',\mu,T)\ \partial_{\mu'}\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n;\mu',0)\,d\mu' \\ &= \frac{1}{T}\int_{-\infty}^{+\infty}(1-f(\mu',\mu,T))f(\mu',\mu,T)\,\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n;\mu',0)\,d\mu' \end{aligned} \tag{5.57}$$

where the last equality involves partial integration and assumes $\sigma(\mu\to-\infty)=0$.

The effect of having a finite temperature is to probe the nonlinear conductivity at different values of the chemical potential. Eq. 5.57 is a kind of weighted average, given by a distribution that is centered and peaked at the Fermi level and has a width of the order k_BT.

We conclude by making some observations on systems with electron-hole symmetry, for which the chemical potential dependence of σ_i^α can be specially simple.

Focusing on the one-photon contributions, the integrand of $\mathcal{I}_j^\alpha(\omega_1)$, at $T=0$ and assuming electron-hole symmetry, is proportional to,

$$\Delta f_{\mathbf{k}vc} = \Theta(\mu-\epsilon_{\mathbf{k}v}) - \Theta(\mu-\epsilon_{\mathbf{k}c}) = \Theta(\Delta\epsilon_{\mathbf{k}cv}-2|\mu|) \to \Theta(\hbar\omega_1-2|\mu|) \tag{5.58}$$

where the last step uses the Dirac delta in Eq. 5.45. The final Heaviside step function is independent of $\mathbf{k}$ and can be pulled out of the $\mathcal{I}_j^\alpha(\omega_1)$ integrals: we have therefore an universal functional dependence on the chemical potential, with frequency dependent coefficients set by the material, the inverse situation to what happens in σ_F. For gapped systems, we could more generally write $\Theta(\hbar\omega_1-\max(2|\mu|,\Delta))$, where $\max(2|\mu|,\Delta)$ is the "effective gap".

Similarly, for the i-photon contribution,

$$\Delta f_{\mathbf{k}vc} \to \Theta(\hbar\omega_1+\cdots+\hbar\omega_i-\max(2|\mu|,\Delta)) \tag{5.59}$$

If the system also respects time-reversal symmetry (Section 6.6), then these terms determine the real part. *The real part of the nonlinear conductivity of a time-reversal and electron-hole symmetric system is always given in terms of step functions in the chemical potential*, aside from the Fermi surface contributions that are present in doped systems at low frequencies. This is exemplified by the expression in Eq. 5.1.

These abrupt changes in the conductivity are due to Pauli blocking (Fig. 4.4a). Depending on the dispersion of the factors multiplying the Heaviside step functions, the steps can be found to vary significantly in magnitude and even be absent or infinite.

On the other hand, the imaginary part is obtained by applying a Hilbert transform to $\mathcal{I}_j^\alpha$ and different systems will have different functional dependencies on the chemical potential. No universal behaviour can be predicted, in this context, for the imaginary part.

The corresponding finite temperature result follows from Eqs. 5.58 and 5.56,

$$\int_{-\infty}^{+\infty} f(\mu', \mu, T)\, \partial_{\mu'} \Theta(\hbar\omega_1 - 2|\mu'|)\, d\mu' = f(-\hbar\omega_1/2, \mu, T) - f(\hbar\omega_1/2, \mu, T) \quad (5.60)$$

or, more generally,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} f(\mu', \mu, T)\, \partial_{\mu'} \Theta(\hbar\omega_1 + \cdots + \hbar\omega_i - \max(2|\mu'|, \Delta))\, d\mu' \\
&= \Theta(\hbar\omega_1 + \cdots + \hbar\omega_i - \Delta) \int_{-\infty}^{+\infty} f(\mu', \mu, T)\, \partial_{\mu'} \Theta(\hbar\omega_1 + \cdots + \hbar\omega_i - 2|\mu'|)\, d\mu' \\
&= \left(f\left(-\frac{\hbar\omega_1 + \cdots + \hbar\omega_i}{2}, \mu, T \right) - f\left(\frac{\hbar\omega_1 + \cdots + \hbar\omega_i}{2}, \mu, T \right) \right) \Theta(\hbar\omega_1 + \cdots + \hbar\omega_i - \Delta)
\end{aligned} \quad (5.61)$$

Introducing a finite temperature smooths the step functions observed when varying the chemical potential. Similar effects manifest in the Hilbert transforms $\mathcal{H}_j^\alpha$, where the peaks tend to broaden and diminish in magnitude.

An Heaviside step function holds out in Eq. 5.61, even at finite temperature: if the Fermi level lies in the band gap, jump discontinuities are still observed due to the gap in the density of states.

5.5 Harmonic generation in monolayer graphene (analytical)

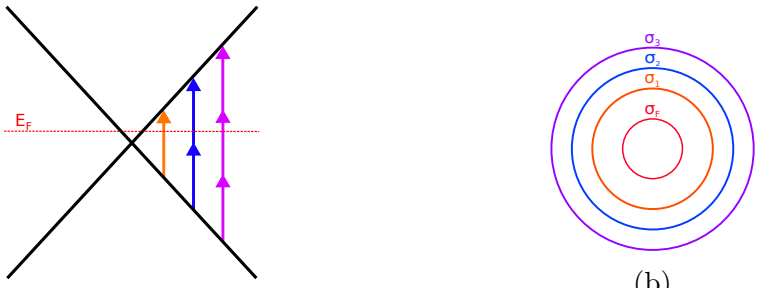
The resonance-based analysis proposed in this chapter, built on the length gauge perturbation theory of Aversa and Sipe [49] and examined in the context of a generic two-band model, aimed at simultaneously clarifying the underlying formal structure of the nonlinear conductivity and streamlining derivations of the conductivities for specific systems. This is best demonstrated by working through an example in detail.

For this purpose, the nonlinear optical response of monolayer graphene, computed numerically in Section 4.6 with the minimal coupling method, is revisited and recalculated with our new tools. This time, analytical expressions will be derived and a low energy effective Hamiltonian, the Dirac Hamiltonian (with or without a mass term), will be adopted, instead of the complete tight binding Hamiltonian demanded by the minimal coupling method.

As always, knowledge of the band structure $\epsilon_{\mathbf{k}a}$ and the non-abelian Berry connection $\mathcal{A}^\beta_{\mathbf{k}ab}$ is required. These were presented in Sections 3.3 and 3.4, more specifically in Eqs. 3.73-3.78, where the possibility of a band gap is included. These fundamental objects often appear differentiated in the integrands of the Π_j^α integrals. Evaluation of the derivatives is a straightforward, if sometimes tedious task (at least by the point one is dealing with second order generalized derivatives of $\mathcal{A}$).

Having garnered the necessary ingredients, the Π_j^α integrals are to be evaluated. Our attention focuses on the relaxation-free limit, at $T = 0$K, where analytical expressions are more accessible. In this limit, the Π_j^α integrals are separated into the $\mathcal{I}_j^\alpha$ integrals and their Hilbert transforms $\mathcal{H}_j^\alpha$ (Eqs. 5.45-5.47).

The $\mathcal{I}_j^\alpha$ integrals, whose integrands contain Dirac deltas, are evaluated by performing the integration along the regions of the FBZ where the respective resonance



(a) (b)

Figure 5.1: (a) Cross section of the Dirac cone with the dark lines showing the linear energy-momentum relation in Eq. 3.61. The arrows stand for incident photons, all of which are assumed to have the same energy (the color scheme serves only to differentiate the different contributions), but probe different regions of the FBZ depending on the number of photons involved. (b) Contours in the FBZ where the resonance conditions that define the different contributions in Eq. 5.4 are met. For the two-dimensional crystal of monolayer graphene, they consist of circles centered at the Dirac point. The red circle is the Fermi surface.

condition is met. For two-dimensional crystals, these usually form closed contours in the FBZ and for the particularly simple, isotropic, band structure of graphene, they consist of circles centered at the Dirac points, as depicted in Fig. 5.1b. In Fig. 5.1, the linear dispersion of graphene was considered, but the contours remain a series of concentric circles for the parabolic dispersion of gapped graphene (Fig. 3.3), with the single difference that $2|\mu|$ has to exceed the band gap for there to be a Fermi surface.

Following the evaluation of the $\mathcal{I}_j^\alpha$ integrals, their Hilbert transforms $\mathcal{H}_j^\alpha$ are obtained (Eq. H.40). More often than not, these cannot be expressed in terms of elementary functions and one has to be content with a numerical computation of these quantities. A greater concern, though, is that, *for an effective model, the Hilbert transforms do not necessarily converge*. Unlike the $\mathcal{I}_j^\alpha$ integrals, which are described by restricted regions of the FBZ, the $\mathcal{H}_j^\alpha$ in Eq. 5.46 may require consideration of the electronic properties of the system over the entire FBZ. If the complete band structure is taken into account (e.g. by using the full tight binding model Hamiltonian), the Hilbert transform in Eq. H.40 runs over the limited bandwidth of the system and converges[9]. It may happen that the integrand decays sufficiently fast for higher energies that it can be approximated by a low energy effective model, but there are no guaranties this will be the case[10].

This is not as in the case of the minimal coupling method, where the difficulties with handling low energy Hamiltonians reflect an deficiency of the method, but is an intrinsic physical limitation on the use of these models. It is perhaps not sufficiently appreciated that continuum Hamiltonians are generally insufficient to retrieve the imaginary part of a conductivity.

Fortunately, for graphene, the Hilbert transforms do converge, provide an appropriate description for optical frequencies near the Dirac point and can even be expressed analytically.

[9] Almost always. Exceptional situations do exist, where this is not the case. See the discussion on undoped gapless graphene at the end of Section 5.5.1, for example.

[10] For an example of a model where the Hilbert transforms do not converge, see Section 5 of [91].

A last point of order. When adopting the Dirac Hamiltonian in graphene, it is important take into account the existence of two Dirac cones (per spin). The integrations on $\mathcal{I}_j^\alpha$ cover the neighbourhoods of both Dirac points, $\mathbf{K}$ and $\mathbf{K}'$, and both the Hamiltonian in Eq. 3.60 and in Eq. 3.65 (or their gapped analogues) are relevant. Instead of doubling the computational effort, the easy way to include both Dirac cones in the calculation is to notice the time-reversal symmetry that relates them (Section 6.6). Graphene possesses time-reversal symmetry and the integrands in Π_j^α can be rewritten in an way that combines the contributions from $\mathbf{k}$ and $-\mathbf{k}$ (see Section 6.6). For a reproduction of the formulae in Section 5.2 with the additional assumption of the time-reversal symmetry, see [91].

The integral evaluation is left to an appendix (Appendix H), where all the $\mathcal{I}_j^\alpha$ and $\mathcal{H}_j^\alpha$ integrals are displayed for monolayer graphene, from linear to third order. Here, the attention will be focused on the inspection of the final expressions for the nonlinear conductivity and its resonance-based decomposition (Eq. 5.4), starting with the gapless system, for which the formulas will be simpler.

5.5.1 $\Delta = 0$

The linear conductivity of graphene is computed by direct application of the formulas of Section 5.2.1 to the dispersion relation in Eq. 3.61 and the Berry connection in Eqs. 3.63 and 3.64. The evaluation of the necessary integrals (F_A^{xx}, $\mathcal{I}_1^{xx}(\omega)$ and $\mathcal{H}_1^{xx}(\omega)$) is done in Eqs. H.13, H.5 and H.6, on which the band gap was set to zero ($\Delta = 0$).

The optical conductivity separates into a Fermi surface contribution and an one-photon contribution:

$$\sigma^{xx}(\omega) = \sigma_F^{xx}(\omega) + \sigma_1^{xx}(\omega) \tag{5.62}$$

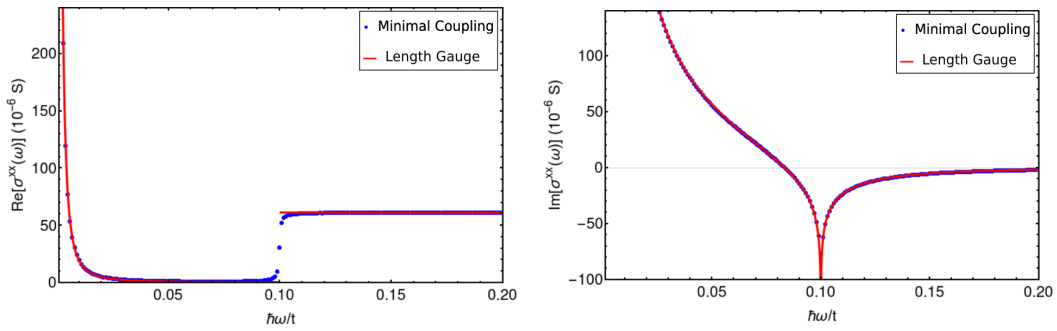
with

$$\sigma_F^{xx}(\omega) = \frac{4\,i\,\sigma_0}{\pi}\,\frac{|\mu|}{\hbar\omega} \tag{5.63}$$

$$\sigma_1^{xx}(\omega) = \sigma_0\left(\Theta\left(\hbar|\omega| - 2|\mu|\right) + \frac{i}{\pi}\ln\left|\frac{\hbar\omega - 2|\mu|}{\hbar\omega + 2|\mu|}\right|\right) \tag{5.64}$$

The Fermi surface contribution is proportional to the chemical potential (and therefore grows with the square root of the carrier concentration $\sigma_F \propto \sqrt{n_e}$) and falls off with the inverse power of the optical frequency. The dispersion of the real part of the one-photon contribution is governed by an Heaviside theta function: it is zero for optical frequencies below the effective gap $2|\mu|$ (Pauli blocking) and suddenly jumps, at $\omega = 2|\mu|$, to a constant value, once interband transitions are allowed. For the imaginary part, the one-photon contribution has a logarithm, that diverges at the effective gap. For $|\omega| \gg 2|\mu|$, the imaginary part becomes negligible and the universal conductivity of undoped graphene is recovered: $\sigma^{xx}(\omega) = \sigma_0 = e^2/4\hbar$ [70].

In Fig. 5.2, the optical conductivity in Eqs. 5.62-5.64, obtained via the length gauge method, is superposed on the linear response computed numerically using the minimal coupling method, previously represented in Fig. 4.2. Excellent agreement is found between numerical and analytical results. The methodologies that are used to arrive at the optical conductivity of graphene in Chapters 4 and 5 are markedly



(a) (b)

Figure 5.2: Real (a) and imaginary (b) parts of the linear optical conductivity of monolayer graphene as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point. The blue dots were numerically computed at those specific frequencies with the minimal coupling method, using a tight binding model and following the algorithm delineated in Chapter 4. The same points when joined produced Fig. 4.2. The red curves do *not* represent a fit, but the result of an independent analytical calculation with the length gauge method, in the relaxation-free limit, by evaluating the integrals in Section 5.2.1. In the real part, the Fermi surface contribution was evaluated for $\gamma = 5 \times 10^{-4}\, t$.

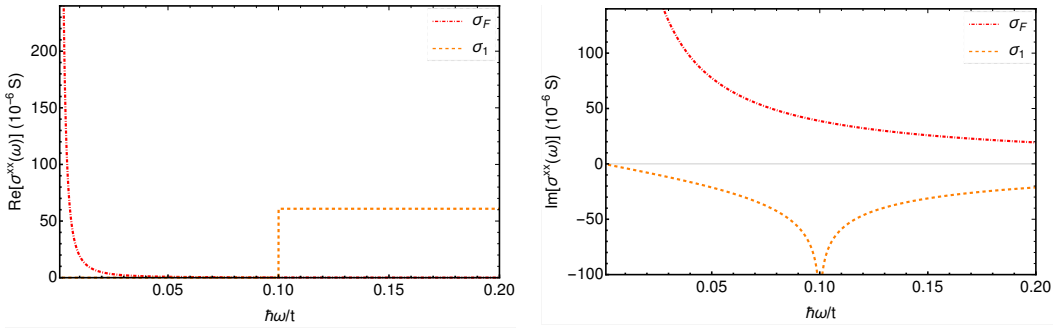
different, despite their equivalence, and having them reproduce the same results is certainly reassuring.

In all honesty, a minor change had to be made to Eq. 5.63 before plotting, or there would be a noticeably difference between the two results. Notice that in the length gauge method used here, the results were considered in the relaxation-free limit, in order to derive analytical expressions. For the minimal coupling method of Chapter 4, the calculations were made with a very small, but finite, value for the relaxation parameter γ.

This is a significant distinction between the two calculations. The slight discrepancies that are observable between the two results, such as the jump in the real part between zero and σ_0 not being instantaneous in the minimal coupling method, can be attributed to this finite value for γ (as in Section 4.6, $\gamma = 0.5\%$ of the effective gap in the minimal coupling computation).

This distinction is most obvious when attending to the real part of the Fermi surface contribution. In Eq. 5.63, σ_F is purely imaginary. This result came from the assumption that $\omega \neq 0$ and the neglect of a Dirac delta that, in the relaxation-free limit, hides at zero frequency: $1/\bar{\omega} \to 1/\omega - \pi\, i\, \delta(\omega)$. Whenever a finite γ is considered, as in the minimal coupling method, the Drude peak will have an width of comparable magnitude, $\sim \gamma$, and will be visible in the dispersion curves. For $\gamma = 0$, as in the analytical calculations, it will not.

It turns out that this difficulty is easily surpassed. Even though performing analytical calculations for finite γ is not usually possible, as it would involve evaluating the Π_j^α integrals directly, instead of $\mathcal{I}_j^\alpha$ and its Hilbert transform, it can done for Fermi surface contributions without issue. Since the F_X^α do not depend on the optical frequency, the analytical calculation of σ_F never required taking the relaxation-free limit. For this reason, it is enough to add $i\,\gamma$ to the optical frequency in Eq. 5.63. Indeed, before representing the length gauge results in Fig. 5.2, the real part of



(a) (b)

Figure 5.3: Resonance-based decomposition of the real (a) and imaginary (b) parts of the linear optical conductivity of monolayer graphene as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point. The Fermi-surface contribution is represented in red and the one-photon contribution in orange, in line with the color coding of Fig. 5.1. In the real part, the Fermi surface contribution is made visible by considering a finite γ.

Eq. 5.63 was replaced by

$$\mathrm{Re}\{\sigma_F^{xx}(\bar{\omega})\} = \frac{4\,\sigma_0}{\pi}\,|\mu|\,\frac{\gamma}{(\hbar\omega)^2+\gamma^2} \tag{5.65}$$

giving the excellent agreement found in Fig. 5.2a at low frequencies.

The imaginary part of σ_F is left unchanged, since, as long as $\gamma \ll \omega$, the changes induced by a finite γ are negligible.

The resonance-based decomposition of the optical conductivity is illustrated in Fig. 5.3, where it is possible to recognize that the Fermi surface contribution dominates the low frequency behavior. For optical frequencies below $2|\mu|$, the real part of the one-photon contribution is identically zero, while the imaginary part tends to zero in the DC limit, leaving only the Fermi surface contribution present. The one-photon contribution is characterized by a step-like increase in the real part, when the optical frequency matches the effective gap, and a negative divergence in the imaginary part that we now recognize as logarithmic.

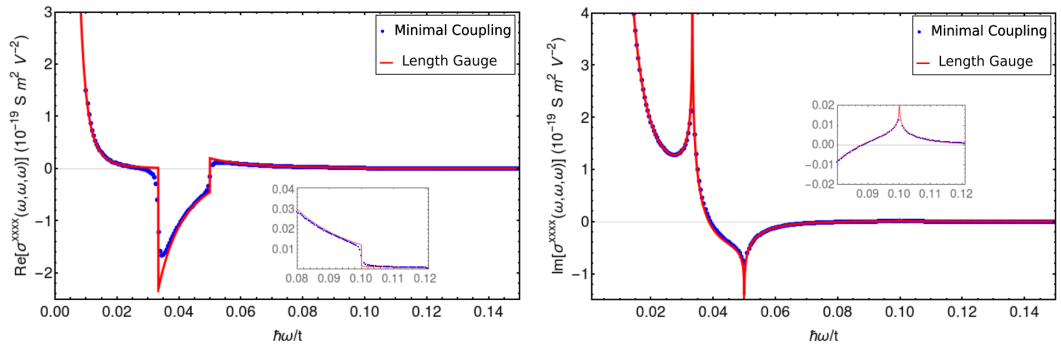
In fact, the case of graphene showcases a general characteristic of the optical response that stems from the Hilbert transforms relating $\mathcal{I}_j^\alpha$ and $\mathcal{H}_j^\alpha$: *discontinuities in the real part* (as those introduced by Heaviside theta functions) *translate into logarithmic divergences in the imaginary part* and vice-versa (see Eq. H.40).

Next, the same analysis is applied to the nonlinear optical response. Monolayer graphene is centrosymmetric, so it has no second order response. The third order optical response is obtained from the formulae in Section 5.2.3, with the necessary integral evaluation performed in Eqs. H.15-H.39 from Appendix H.

In a resonance-based analysis, the third order optical conductivity is decomposed into Fermi surface, one-, two- and three-photon contributions,

$$\sigma^{xxxx}(\omega,\omega,\omega) = \sigma_F^{xxxx}(\omega,\omega,\omega) + \sigma_1^{xxxx}(\omega,\omega,\omega) + \sigma_2^{xxxx}(\omega,\omega,\omega) + \sigma_3^{xxxx}(\omega,\omega,\omega) \tag{5.66}$$

with



(a) (b)

Figure 5.4: Real (a) and imaginary (b) parts of the third order optical conductivity of monolayer graphene as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point. The blue dots were numerically computed at those specific frequencies with the minimal coupling method, using a tight binding model and following the algorithm delineated in Chapter 4. The same points when joined produced Fig. 4.8. The red curves do *not* represent a fit, but the result of an independent analytical calculation with the length gauge method, in the relaxation-free limit, by evaluating the integrals in Section 5.2.3. In the real part, the Fermi surface contribution was evaluated for $\gamma = 5 \times 10^{-4}\, t$. In the inset, it is possible to discern the smaller one-photon feature.

$$\sigma_F^{xxxx}(\omega,\omega,\omega) = +\frac{24\, i\, C_0}{\pi|\mu|(\hbar\omega)^3} \tag{5.67}$$

$$\sigma_1^{xxxx}(\omega,\omega,\omega) = -\frac{17\, C_0}{(\hbar\omega)^4}\left(\Theta\left(\hbar|\omega| - 2|\mu|\right) + \frac{i}{\pi}\ln\left|\frac{\hbar\omega - 2|\mu|}{\hbar\omega + 2|\mu|}\right|\right) + \frac{71\, i\, C_0}{\pi|\mu|(\hbar\omega)^3} \tag{5.68}$$

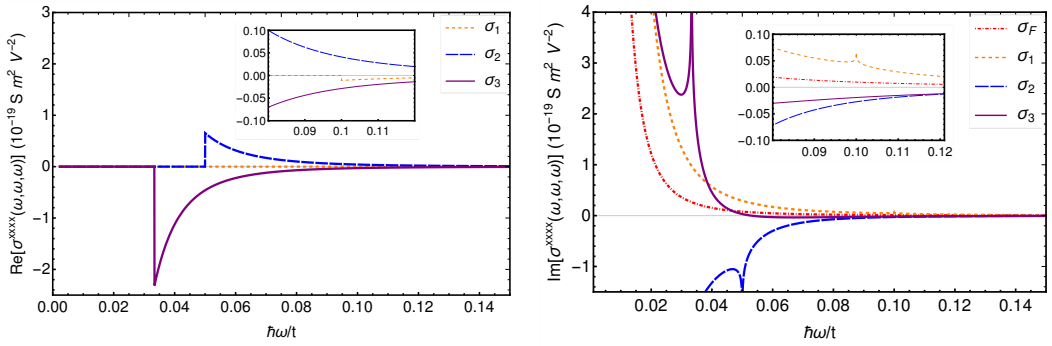
$$\sigma_2^{xxxx}(\omega,\omega,\omega) = +\frac{64\, C_0}{(\hbar\omega)^4}\left(\Theta\left(2\hbar|\omega| - 2|\mu|\right) + \frac{i}{\pi}\ln\left|\frac{2\hbar\omega - 2|\mu|}{2\hbar\omega + 2|\mu|}\right|\right) - \frac{32\, i\, C_0}{\pi|\mu|(\hbar\omega)^3} \tag{5.69}$$

$$\sigma_3^{xxxx}(\omega,\omega,\omega) = -\frac{45\, C_0}{(\hbar\omega)^4}\left(\Theta\left(3\hbar|\omega| - 2|\mu|\right) + \frac{i}{\pi}\ln\left|\frac{3\hbar\omega - 2|\mu|}{3\hbar\omega + 2|\mu|}\right|\right) - \frac{63\, i\, C_0}{\pi|\mu|(\hbar\omega)^3} \tag{5.70}$$

Similarly to linear order, the real part of the third order conductivity is marked by jump discontinuities and the imaginary part by the associated logarithmic divergences, whenever either one-, two- or three-photon frequencies match the effective gap. The real parts of the one-, two- and three-photon contributions are zero when the respective optical frequencies lie below the gap, but this time the imaginary parts do not tend to zero in the DC limit (in fact, they diverge!).

Unlike linear order, the magnitude of the steps and logarithms falls off with ω^{-4}, with the exception of an additional set of terms that decay less rapidly with frequency ($\sim \omega^{-3}$) and have a curious inverse dependence on the chemical potential. The Fermi surface contribution consists of a term of this type.

Again, the analytical results are compared with those obtained numerically in the previous chapter (Fig. 4.8). The real and imaginary parts are displayed in Fig. 5.4, where the minimal coupling and length gauge methods produce identical curves.



(a) (b)

Figure 5.5: Resonance-based decomposition of the real (a) and imaginary (b) parts of the third order optical conductivity of monolayer graphene as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point, in the relaxation-free limit ($\gamma = 0$). The color coding is the same as in Fig. 5.1.

Once more, to obtain the perfect fit seen in the real part at low frequencies, the formula for the Fermi surface contribution had to be extended for finite γ, simply replacing $\omega \to \bar{\omega}$. At higher frequencies, it is possible to discern minor differences between the numerical and analytical evaluations near the sharper features of the real part in Fig. 5.4a. This is to be expected, and these differences become all the less relevant, the smaller the value adopted for γ in the numerical computations. As $\gamma \to 0$, the line traced by the blue dots (minimal coupling method) grows closer to the idealized red line (length gauge method).

The resonance-based decomposition of the third order conductivity is analyzed in Fig. 5.5. The expected steps are observed in the real part of the one-, two- and three photon contributions. It was suggested in [76, 78] that the small magnitude of the one-photon feature in Fig. 5.4 is due to the cancellation between different contributions of alternating signs (the coefficients of the Heaviside theta functions produce $-17+64-45 = 2$). Although this serves as a partial explanation, the truth is that the one-photon contribution in Fig. 5.5 is already small, prior to taking the other contributions into account. The minuteness of the one-photon feature in the third order contribution[11] is mostly because of fast decrease of the response with the optical frequency $\mathrm{Re}\{\sigma^{xxxx}\} \sim \omega^{-4}$.

The imaginary parts display the expected logarithmic divergences, partners of the steps in the real part, and have a somewhat unexpected divergence in the DC limit. This is not too surprising, since this is the doped system and there is a concentration of free charge carriers that dominate the response at low frequencies, but one could intuitively assign this part of the response to the Fermi surface contribution. As it turns out, *in the imaginary part, Fermi surface contributions seem to be insufficient to provide a complete description of the nonlinear optical response at low frequencies.* This puts into question the validity of Boltzmann equation methods in describing the imaginary part of the optical response, even in the regime where the optical frequencies lie significantly below the effective gap.

An advantage of decomposing the nonlinear conductivity in this manner is that

[11] For third harmonic generation. The situation is inverted in the case of difference frequency generation [78].

it identifies the components of the theory that are relevant in any given region of the spectrum. If the parameters of an experiment on third harmonic generation dictate that a narrow window containing the two-photon feature is relevant, the theory may approximated for this region by taking only the two-photon contribution into account. This would be simpler to compute, since the two-photon contribution contains only Π_j^α integrals with $j = 1, 3$ or 5 (Eq. 5.33).

A curious special case is that of undoped graphene. A hasty look through the expressions for the third order conductivity might lead one to think that it diverges when the chemical potential is set to zero. But when summing all the contributions (Eqs. 5.67-5.70) to the third order conductivity the different factors of μ^{-1} cancel (Eq. 5.1).

Nonetheless, there have been some concerns with this divergence in the literature. In the work of Cheng et al. [77] a divergence of such nature was found. But, as pointed out by the authors, it appears only when introducing two distinct damping parameters, one for intraband and one for interband resonances. This suggests that such a divergence is an artifact of the phenomenological description of relaxation that was adopted in their work. In the method used here, there is a single parameter γ describing relaxation, but this parameter might be made different at low and high frequencies (as pointed out in Section 2.5, any even function $\gamma = \gamma(\omega)$ will do) and still no divergences for $\mu = 0$ are found.

Similarly, for the Boltzmann equation approach adopted in the pioneering works of [72, 74], the third order conductivity was found to diverge when $\mu \to 0$. The reason is transparent in our framework: the Boltzmann equation approach takes only the contribution from the Fermi surface into account (purely intraband transitions) and this one can be seen from Eq. 5.67 to indeed diverge for a zero chemical potential. Boltzmann equation methods are usually considered appropriate to describe the low frequency dispersion of the nonlinear conductivity as long as there is a gap, either an actual band gap or an effective one set by Pauli blocking, below which all contributions other than the ones defined by the Fermi surface become negligible. For undoped graphene, however, this is never the case and the remaining contributions in Eqs. 5.68-5.70 must always be considered.

Unfortunately, a real difficulty can still be found for undoped systems with zero gap in the DC limit. Even though the $\mu = 0$ third order conductivity of graphene is well defined for any finite frequency, it shows a frequency dependence ω^{-4} (see Eq. 2 of [76]) diverging when $\omega \to 0^+$. As seen before, in Section 4.6, this result holds even for a tight binding model defined over the entire FBZ and is presumably due to a breakdown in perturbation theory.

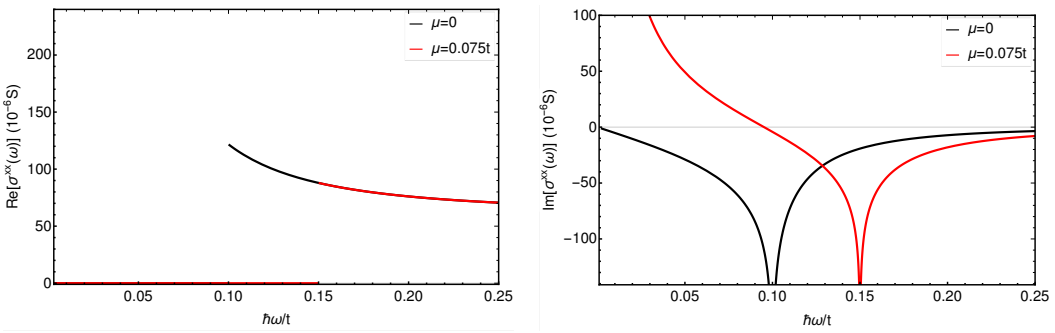
5.5.2 $\Delta \neq 0$

Consider now that a mass term is inserted in the Dirac Hamiltonian (Eq. 3.72), generating a band gap ($\Delta \neq 0$). Once more, the optical response is derived, analytically, with the formulae of Section 5.2. The evaluation of the integrals is done in Appendix H.

The end result is that the linear optical conductivity gives

$$\sigma^{xx}(\omega) = \sigma_F^{xx}(\omega) + \sigma_1^{xx}(\omega) \tag{5.71}$$

with



(a) (b)

Figure 5.6: Real (a) and imaginary (b) parts of the linear optical conductivity of gapped graphene, with $\Delta = 0.1\,t$, as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point. The response was obtained analytically (Eqs. 5.71-5.73) in the relaxation-free limit ($\gamma = 0$). The black and red curves refer to clean ($\mu = 0$) and doped ($\mu = 0.075\,t$) systems, corresponding to effective gaps of $0.1\,t$ and $0.15\,t$, respectively.

$$\sigma_F^{xx}(\omega) = \frac{4\,i\,\sigma_0}{\pi}\,\frac{|\mu|}{\hbar\omega}\left(1-\left(\frac{\Delta}{2|\mu|}\right)^2\right)\Theta\left(2|\mu|-\Delta\right) \tag{5.72}$$

$$\sigma_1^{xx}(\omega) = \sigma_0\left(1+\left(\frac{\Delta}{\hbar\omega}\right)^2\right)\left(\Theta\left(\hbar|\omega|-\max(\Delta,2|\mu|)\right)+\frac{i}{\pi}\ln\left|\frac{\hbar\omega-\max(\Delta,2|\mu|)}{\hbar\omega+\max(\Delta,2|\mu|)}\right|\right) + \frac{2\,i\,\sigma_0}{\pi}\,\frac{\Delta^2}{\hbar\omega\,\max(\Delta,2|\mu|)} \tag{5.73}$$

As before, the real part of the one-photon contribution is proportional to an Heaviside theta function and the imaginary part contains a logarithm. However, the step and the logarithmic divergence appear now at the effective gap,

$$\Delta_{eff} \equiv \max(\Delta, 2|\mu|) \tag{5.74}$$

The magnitude of these features is dependent on a series of terms in even powers of Δ/ω. For the Fermi surface contribution, the chemical potential is the relevant energy scale and the magnitude of this contribution is determined by a series of terms in even powers of Δ/μ. Naturally, when we set $\Delta = 0$, only the zeroth order terms remain and the Fermi surface and one-photon contributions reproduce the previous results in Eqs. 5.63 and 5.64, respectively.

In addition to this, there is a curious term in the one-photon contribution that is proportional to the adimensional ratio $\Delta^2/\omega\,\Delta_{eff}$. It vanishes in the absence of a band gap, it varies inversely with frequency and for a doped system, $2|\mu| > \Delta$, also goes with the inverse of μ. If the Fermi level falls below the band gap, $\Delta_{eff} = \Delta$ and the term is independent of the chemical potential (there are no worries of divergences with μ^{-1} in gapped graphene).

The conductivity in Eqs. 5.71-5.73 is depicted in Fig. 5.6, for two scenarios: $2|\mu| < \Delta$, which implies $\Delta_{eff} = \Delta$, and $2|\mu| > \Delta$, for which $\Delta_{eff} = 2|\mu|$.

In the undoped system, the real part has a step at the band gap and the imaginary

part a logarithmic divergence, as had already been seen in Fig. 4.1. The real part vanishes below the band gap and the imaginary part tends to zero in the DC limit.

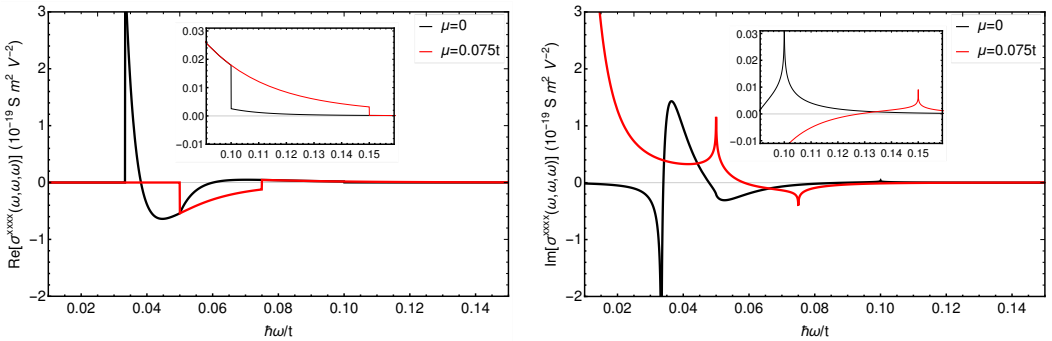
If the Fermi level is raised above the band gap, a Drude peak appears at low frequencies. Since no comparisons are being made with results from the minimal coupling method, all contributions, including σ_F, are evaluated in the relaxation-free limit. As a consequence, the Drude peak is not visible in the real part (Fig. 5.6a), consisting of a Dirac delta of infinitesimal width. It is, however, clearly visible in the imaginary part. The other, predictable, change is that features shift from the band gap to the effective gap at $2|\mu|$. All the chemical potential dependence of the real part of the conductivity is in the Heaviside step function, which means that increasing the doping level will affect the dispersion of the real part only by nullifying the optical response below the Fermi level. This is a general point that has been emphasized in Section 5.4 and is manifest in Fig. 5.6a.

Ideally, the discussion would now advance to the second order optical conductivity of gapped graphene. Since inversion symmetry is broken with the introduction of different on-site energies, the second order optical response is nonzero, as demonstrated in Section 4.6.3. Unfortunately, this second order response cannot be captured by the Dirac Hamiltonian used here. If the equations and integrals of Section 5.2.2 are evaluated, even with a mass term in the Hamiltonian, the second order conductivity returns zero. It is necessary to use a more accurate effective model, by carrying the low energy expansion (Eq. 3.58) to the next order, to be able to properly capture the second order conductivity. For more on this topic, consult the work in [84].

Knowing this, we proceed straight to the third order optical conductivity of gapped graphene:

$$\sigma^{xxxx}(\omega,\omega,\omega) = \sigma_F^{xxxx}(\omega,\omega,\omega) + \sigma_1^{xxxx}(\omega,\omega,\omega) + \sigma_2^{xxxx}(\omega,\omega,\omega) + \sigma_3^{xxxx}(\omega,\omega,\omega) \tag{5.75}$$

with



(a) (b)

Figure 5.7: Real (a) and imaginary (b) parts of the third order optical conductivity of gapped graphene, with $\Delta = 0.1\,t$, as a function of the optical frequency, normalized to the tight binding parameter t, near the Dirac point. The response was obtained analytically (Eqs. 5.75-5.79) in the relaxation-free limit ($\gamma = 0$). The black and red curves refer to clean ($\mu = 0$) and doped ($\mu = 0.075\,t$) systems, corresponding to effective gaps of $0.1\,t$ and $0.15\,t$, respectively.

$$\sigma_F^{xxxx}(\omega,\omega,\omega) = \frac{24\,i\,C_0}{\pi\,|\mu|\,(\hbar\omega)^3}\left(1+2\left(\frac{\Delta}{2\,|\mu|}\right)^2-3\left(\frac{\Delta}{2|\mu|}\right)^4\right)\Theta(2|\mu|-\Delta) \tag{5.76}$$

$$\begin{aligned}\sigma_1^{xxxx}(\omega,\omega,\omega) &= \frac{2\,i\,C_0}{\pi\,(\hbar\omega)^3\,\Delta_{eff}}\left(71-22\left(\frac{\Delta}{\hbar\omega}\right)^2+15\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\\ &+\frac{2\,i\,C_0}{3\,\pi\,\hbar\omega\,\Delta_{eff}^3}\left(154\left(\frac{\Delta}{\hbar\omega}\right)^2+15\left(\frac{\Delta}{\hbar\omega}\right)^4\right)-\frac{498\,i\,C_0\,\Delta^4}{5\,\pi\,(\hbar\omega)^3\,\Delta_{eff}^5}\\ &+\frac{C_0}{(\hbar\omega)^4}\left(-17-22\left(\frac{\Delta}{\hbar\omega}\right)^2+15\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\left(\Theta\left(\hbar|\omega|-\Delta_{eff}\right)+\frac{i}{\pi}\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|\right)\end{aligned} \tag{5.77}$$

$$\begin{aligned}\sigma_2^{xxxx}(\omega,\omega,\omega) &= \frac{2\,i\,C_0}{\pi\,(\hbar\omega)^3\,\Delta_{eff}}\left(-32+64\left(\frac{\Delta}{\hbar\omega}\right)^2-24\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\\ &-\frac{64\,i\,C_0}{3\,\pi\,\hbar\omega\,\Delta_{eff}^3}\left(2\left(\frac{\Delta}{\hbar\omega}\right)^2+3\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{192\,i\,C_0\,\Delta^4}{5\,\pi\,(\hbar\omega)^3\,\Delta_{eff}^5}\\ &+\frac{C_0}{(\hbar\omega)^4}\left(64+32\left(\frac{\Delta}{\hbar\omega}\right)^2-12\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\left(\Theta\left(2\hbar|\omega|-\Delta_{eff}\right)+\frac{i}{\pi}\ln\left|\frac{2\hbar\omega-\Delta_{eff}}{2\hbar\omega+\Delta_{eff}}\right|\right)\end{aligned} \tag{5.78}$$

$$\begin{aligned}\sigma_3^{xxxx}(\omega,\omega,\omega) &= \frac{2\,i\,C_0}{\pi\,(\hbar\omega)^3\,\Delta_{eff}}\left(-63-42\left(\frac{\Delta}{\hbar\omega}\right)^2+9\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\\ &+\frac{2\,i\,C_0}{\pi\,(\hbar\omega)\,\Delta_{eff}^3}\left(-78\left(\frac{\Delta}{\hbar\omega}\right)^2+27\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{1026\,i\,C_0\,\Delta^4}{5\,\pi\,(\hbar\omega)^3\,\Delta_{eff}^5}\\ &+\frac{C_0}{(\hbar\omega)^4}\left(-45-14\left(\frac{\Delta}{\hbar\omega}\right)^2+3\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\left(\Theta\left(3\hbar|\omega|-\Delta_{eff}\right)+\frac{i}{\pi}\ln\left|\frac{3\hbar\omega-\Delta_{eff}}{3\hbar\omega+\Delta_{eff}}\right|\right)\end{aligned} \tag{5.79}$$

These expressions are longer, but they form a generalization of Eqs. 5.67-5.70 that is analogous to that observed in linear order. The band gap is again replaced by the effective gap in the steps and logarithmic divergences that characterize the real and imaginary parts, respectively. Again, the magnitude of the contributions is defined by a series of terms in even powers of Δ/ω, or $\Delta/|\mu|$ in the case of the Fermi surface contribution. More of those added terms that vary with products of powers of the optical frequency and the effective gap influence the imaginary part of the nonlinear optical response. These terms are themselves multiplied by an sequence of even powers of Δ/ω. It is a more complicated set of formulas, but it is quickly recognized that when $\Delta \to 0$, the third order conductivity of the previous subsection is recovered.

The entire third order conductivity is represented in Fig. 5.7, again for both the doped and undoped systems. In both cases, the steps and divergences alternate signs and diminish in magnitude for higher frequencies. Interestingly, the undoped system has no jump discontinuity in the real part at $2\hbar\omega = \Delta$. Instead, it has what could be termed a "suppressed step" (Fig. 5.7a). This is an excellent opportunity to attest to the usefulness of the resonance-based decomposition, since the absence of a jump discontinuity can be understood, and predicted, by noticing in Eq. 5.78 that $\mathrm{Re}\{\sigma_2^{xxxx}\} = 0$ when $\omega = \Delta/2\hbar$.

Setting the Fermi level above the band gap leads to a shift in the one-, two- and three-photon features and the emergence of a Drude peak at low frequencies, as expected. Even though all the dependence of the real part on the chemical potential still lies in the blocking of interband transitions, it is no longer true that the effect of raising the Fermi level can be immediately read by cutting out part of the dispersion curve, since the optical response is now built out of a combination of the three Heaviside step functions.

Chapter 6

The role of symmetry

After presenting the definition of the nonlinear conductivity and its importance in nonlinear optics, the question that dominated our discussion was: how to efficiently compute it for a crystal, if given some fundamental electronic properties (ϵ and $\mathcal{A}$)? Chapter 4 resolved this question by describing a numerical procedure of wide applicability and Chapter 5 made analytical calculations on effective models as simple as they can be.

We now shift focus to a different question: what kind of functions could serve as the nonlinear conductivity of a crystal? In other words, what constraints does the conductivity obey? In the previous section, much progress was made on this front by presenting its detailed structure. The nonlinear conductivity was shown to always be built from contributions relating to the Fermi surface, the one-photon resonance, two-photon resonance and so on. Universal characteristics of the nonlinear conductivity were identified based on the general form of these contributions.

For the remainder of this thesis, we take this analysis further by taking into account the role of symmetry. In particular, there are certain symmetries that are always obeyed by the nonlinear conductivity, no matter the medium.

The first symmetries to be considered are those that apply to all nonlinear response functions, not just the optical conductivity (Section 6.1). After these are briefly reviewed, most of the discussion is devoted to overall permutation symmetry, typically associated to lossless systems (Section 6.2), but here generalized by extending the analysis into the complex plane (Section 6.3), as in [27].

Overall permutation symmetry is an well-known symmetry in nonlinear optics, but new insight is gained when it is applied to the results of the previous chapter. With use of the resonance-based decomposition of the nonlinear conductivity, overall permutation symmetry permits a clear identification of the dissipative part of the response function (Section 6.4) and the derivation of new, unexpected, connections between optical resonances involving different numbers of photons (Section 6.5). For instance, the two-photon and one-photon contributions are shown to be inextricably linked at second order.

A brief account of time-reversal symmetry closes the chapter (Section 6.6). Once again, when the symmetry is seen from the perspective of the resonance-based analysis of Chapter 5, its mathematical consequences are made abundantly clear. Signatures of broken time-reversal symmetry are identified in the nonlinear optical response, namely a nonvanishing real part for the nonlinear conductivity when optical frequencies lie below the band gap.

6.1 A summary of symmetry

Besides providing insight and helping reduce the space of possibilities on what the dispersion of a nonlinear response function could look like, symmetries are of practical use as well. This is exemplified by two symmetries that were used in previous sections to simplify calculations.

The first was intrinsic permutation symmetry (Eq. 2.32), introduced at the end of Section 2.2, as a direct consequence of the definition of the nonlinear response function. Only the portion of the nonlinear response function that respects intrinsic permutation symmetry is physical, since any remaining contribution vanishes when performing the integration in Eq. 1.9 to produce the observable J.

This is irrelevant in linear order. In second order, it tell us that expressions like those for the second order conductivity $\sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)$ in Eqs. 2.63 and 4.24 must be symmetrized in the arguments $(\{\alpha_1,\omega_1\},\{\alpha_2,\omega_2\})$,

$$\sigma_S^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) \equiv \frac{1}{2}\left(\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)\right) \tag{6.1}$$

in order to ensure that

$$\sigma_S^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_S^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1) \tag{6.2}$$

Similarly in third order,

$$\begin{aligned}\sigma_S^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) &= \sigma_S^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}_1,\bar{\omega}_3,\bar{\omega}_2) = \sigma_S^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}_2,\bar{\omega}_1,\bar{\omega}_3)\\ = \sigma_S^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2,\bar{\omega}_3,\bar{\omega}_1) &= \sigma_S^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3,\bar{\omega}_1,\bar{\omega}_2) = \sigma_S^{\beta\alpha_3\alpha_2\alpha_1}(\bar{\omega}_3,\bar{\omega}_2,\bar{\omega}_1)\end{aligned} \tag{6.3}$$

where the physical part of the third order conductivity, σ_S, is derived from Eq. 2.64 or Eq. 4.25 by symmetrization with respect to $(\{\alpha_1,\omega_1\},\{\alpha_2,\omega_2\},\{\alpha_3,\omega_3\})$.

The other symmetry that has been of use in past sections is the symmetry of the crystal. Any crystal lattice, even if it is not a Bravais lattice, will be invariant under a group of symmetries composed of some subset of discrete translations, rotations, reflections, spatial inversion and their combinations. The study of crystal symmetries is made in the context of group theory and categorization of all possible spatial groups and the tensor relations they impose can be found in standard textbooks on the subject [88].

Due to crystal symmetry, not all tensor components of linear and nonlinear response functions will be independent. Like intrinsic permutation symmetry, crystal symmetry defines general relations valid for any response functions, but these also depend on the character of the vector field, that is, on how the classical fields themselves transform under spatial transformations[1].

In previous calculations of the nonlinear optical response of monolayer graphene, crystal symmetry was used, in Sections 4.6 and 5.5, to reduce the number of tensor components to evaluate. In the resonance-based analysis of the previous chapter, it was discussed (Section 5.2.3) that these tensor symmetries apply not only to the full nonlinear conductivity, but to each Π_j^α that must be evaluated, reducing the

[1]For example, in the case of the nonlinear conductivity, $E(t)$ is a polar vector, changing sign when performing spatial inversion.

number of "Fermi golden rule"-type computations and Hilbert transforms that are necessary.

Another condition that is applicable to any response function defined by Eq. 2.24 is the reality condition. Since both J and E are real quantities, so must be our response function,

$$\left(\sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n)\right)^* = \sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n) \tag{6.4}$$

which, in the frequency domain, translates to

$$\left(\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)\right)^* = \sigma^{\beta\alpha_1\ldots\alpha_n}(-\bar{\omega}_1^*,\ldots,-\bar{\omega}_n^*) \tag{6.5}$$

Finally, a more subtle, but still completely general, symmetry is the one brought by causality. At the simplest level, causality implies that the induced current $J(t)$ must depend only on the values of the optical field $E(t')$ at prior times ($t' < t$). It is straightforward to prove that this statement is equivalent to the following condition on the response function,

$$\sigma^{\beta\alpha}(t) = \sigma^{\beta\alpha}(t)\,\Theta(t) \tag{6.6}$$

at linear order, and more generally

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n) = \sigma^{\beta\alpha_1\ldots\alpha_n}(t_1,\ldots,t_n)\,\Theta(t_1)\ldots\Theta(t_n) \tag{6.7}$$

for a nonlinear response function.

A more restrictive condition could be put on the response functions by pulling on the postulates of special relativity [95], but this is not relevant for our discussions[2].

The challenge is now to derive the restrictions imposed by causality on the frequency domain response functions. The most direct approach to arrive at the Kramers-Krönig relations comes from knowing that the Fourier transform of a Heaviside step function is

$$\int_{-\infty}^{+\infty} \Theta(t)\,e^{i\omega t}\,dt = \pi\,\delta(\omega) + \frac{i}{\omega} \tag{6.8}$$

Applying a Fourier transform to Eq. 6.6,

$$\begin{aligned}\sigma^{\beta\alpha}(\omega) &= \sigma^{\beta\alpha}(\omega) * \left(\pi\,\delta(\omega) + \frac{i}{\omega}\right) \\ &= \frac{\sigma^{\beta\alpha}(\omega)}{2} + i\fint_{-\infty}^{+\infty} \frac{d\omega'}{2\pi}\,\frac{\sigma^{\beta\alpha}(\omega')}{\omega-\omega'}\end{aligned} \tag{6.9}$$

where $*$ stands for a convolution[3]. It results that,

$$\sigma^{\beta\alpha}(\omega) = \frac{1}{\pi\,i}\fint_{-\infty}^{+\infty} \frac{\sigma^{\beta\alpha}(\omega')}{\omega'-\omega}\,d\omega' \tag{6.10}$$

[2]Since we are ignoring spatial dispersion, it does not matter how long it would take for light to travel from a place where the optical field varied to where a change in current would be induced.

[3]As defined in the frequency domain: $f(\omega) * g(\omega) = \int_{-\infty}^{+\infty} f(\omega')\,g(\omega-\omega')\,d\omega'/2\pi$.

Putting it simply, causality implies that the response functions in the frequency domain must be invariant under the application of a Hilbert transform, aside from the multiplicative factor i.

Due to this i factor, Eq. 6.10 provides a link between the real and imaginary parts,

$$\mathrm{Re}\{\sigma^{\beta\alpha}(\omega)\} = +\frac{1}{\pi}\fint_{-\infty}^{+\infty}\frac{\mathrm{Im}\{\sigma^{\beta\alpha}(\omega')\}}{\omega'-\omega}\,d\omega' \tag{6.11}$$

$$\mathrm{Im}\{\sigma^{\beta\alpha}(\omega)\} = -\frac{1}{\pi}\fint_{-\infty}^{+\infty}\frac{\mathrm{Re}\{\sigma^{\beta\alpha}(\omega')\}}{\omega'-\omega}\,d\omega' \tag{6.12}$$

so that one could be derived from another.

These are the linear Kramers-Krönig relations. The previous reasoning can be applied to any argument of a nonlinear response function, yielding

$$\sigma^{\beta\alpha_1\ldots\alpha_i\ldots\alpha_n}(\omega_1,\ldots,\omega_i,\ldots,\omega_n) = \frac{1}{\pi\,i}\fint_{-\infty}^{+\infty}\frac{\sigma^{\beta\alpha_1\ldots\alpha_i\ldots\alpha_n}(\omega_1,\ldots,\omega_i',\ldots,\omega_n)}{\omega_i'-\omega_i}\,d\omega_i' \tag{6.13}$$

the nonlinear Kramers-Krönig relations.

Other versions of the nonlinear Kramers-Kröning relations exist, the most general of which was proposed and proved in [95],

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\omega_1+p_1\,\omega,\ldots,\omega_n+p_n\,\omega) = \frac{1}{\pi\,i}\fint_{-\infty}^{+\infty}\frac{\sigma^{\beta\alpha_1\ldots\alpha_n}(\omega_1+p_1\,\omega',\ldots,\omega_n+p_n\,\omega')}{\omega'-\omega}\,d\omega' \tag{6.14}$$

with $p_i \geq 0$ for all i, provided that at least one $p_i \neq 0$.

The most common use of these relations in optics is to derive refraction from absorption. This is true for linear and nonlinear optics. Absorption is both simpler to measure and to compute, and changes in the refractive index are usually more easily derived from the measurement of absorption coefficients than by direct measurements. By the end of this chapter, it should be clear why the part of the nonlinear conductivity related to absorption (Section 6.4) is more accessible in an analytical calculation.

For many systems (Section 6.6), the real part of the conductivity is related to absorption and the imaginary part to refraction [3]. The passage from absorption to refraction is then made by equations like Eqs. 6.11 and 6.12. But the application of an Hilbert transform to the entire real part of a nonlinear conductivity is no easy task; just consider the complexity of the third order conductivity. It must be performed numerically, unless the system happens to be specially simple.

In the previous chapter, a better way to apply the Hilbert transforms was delineated, one that is amenable to analytical calculations (though, more often than not, the transform must still be performed numerically). This was possible by isolating the Π_j integrals and taking the relaxation-free limit, ending up with the integrals $\mathcal{I}_j$ and $\mathcal{H}_j$, connected by an Hilbert transform. It is worth noting that, from the standpoint of getting an analytical expression for the imaginary part, this is a substantial advancement.

There are two more symmetries that will be addressed in this text: overall permutation symmetry and time-reversal symmetry. With these, we leave the perspective

of the general nonlinear response function defined in Section 2.2 and narrow our focus to the specificity of the nonlinear conductivity.

6.2 Lossless media

There is a symmetry that is characteristic of a lossless medium, here defined as one where the energy transferred to the medium from a light pulse propagating through it is zero:

$$\int_{-\infty}^{+\infty} J^\beta(t)\, E^\beta(t)\, dt = 0 \tag{6.15}$$

In the case of a linear medium,

$$\begin{aligned}
\int_{-\infty}^{+\infty} J^\beta(t)\, E^\beta(t)\, dt &= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, J^\beta(\omega)\, E^\beta(-\omega) \\
&= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \frac{1}{2} \left(J^\beta(\omega)\, E^\beta(-\omega) + J^\beta(-\omega)\, E^\beta(\omega) \right) \\
&= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \frac{1}{2} \left(J^\beta(\omega)\, (E^\beta(\omega))^* + (J^\beta(\omega))^*\, E^\beta(\omega) \right) \\
&= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \mathrm{Re}\{ J^\beta(\omega)\, (E^\beta(\omega))^* \} \\
&= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \mathrm{Re}\{ \sigma^{\beta\alpha}(\omega)\, E^\alpha(\omega)\, (E^\beta(\omega))^* \} \\
&= \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \frac{1}{2} \left(\sigma^{\beta\alpha}(\omega) + (\sigma^{\alpha\beta}(\omega))^* \right) E^\alpha(\omega)\, (E^\beta(\omega))^* \\
&= 0
\end{aligned} \tag{6.16}$$

where vanishing relaxation was assumed: $\sigma^{\beta\alpha}(\omega) \equiv \lim_{\gamma \to 0^+} \sigma^{\beta\alpha}(\bar{\omega})$.

For it to be a property of the medium, Eq. 6.16 must be true for any $E(\omega)$. Hence,

$$\sigma^{\beta\alpha}(\omega) + (\sigma^{\alpha\beta}(\omega))^* = 0 \tag{6.17}$$

which states that the conductivity tensor must be anti-hermitian. This is a symmetry of the optical conductivity, a direct consequence of energy conservation and the requirement of no optical losses.

It will prove insightful to rewrite this relation in terms of the susceptibility, $\chi(\omega) = -i\,\omega\,\sigma(\omega)$,

$$\chi^{\beta\alpha}(\omega) = (\chi^{\alpha\beta}(\omega))^* \tag{6.18}$$

which must be hermitian.

A similar reasoning can be applied to a lossless nonlinear medium. Starting with the second order contribution to the current,

$$\int_{-\infty}^{+\infty} J^{\beta(2)}(t)\, E^\beta(t)\, dt = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, J^{\beta(2)}(\omega)\, E^\beta(-\omega) = 0 \tag{6.19}$$

where we can replace $J^{(2)}$ by the second term in the expansion of Eq. 1.5,

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega_2}{2\pi}\, \sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)\, E^{\alpha_1}(\omega_1)\, E^{\alpha_2}(\omega_2)\, E^{\beta}(-\omega_{12}) \\
&= \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega_2}{2\pi}\, E^{\alpha_1}(\omega_1)\, E^{\alpha_2}(\omega_2)\, E^{\beta}(-\omega_{12}) \\
&\quad \times \frac{1}{3}\left(\sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) + \sigma^{\alpha_2\alpha_1\beta}(\omega_1,-\omega_{12})\right)
\end{aligned} \tag{6.20}$$

with the abbreviation $\omega_{12} \equiv \omega_1 + \omega_2$ and the last step involving changes of variables[4]. Replacing the conductivities in the integrand by susceptibilities,

$$\begin{aligned}
&= -\,i \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi} \int_{-\infty}^{+\infty} \frac{d\omega_2}{2\pi}\, E^{\alpha_1}(\omega_1)\, E^{\alpha_2}(\omega_2)\, E^{\beta}(-\omega_{12}) \\
&\quad \times \frac{1}{3}\left(\omega_{12}\,\chi^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) - \omega_1\,\chi^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) - \omega_2\,\chi^{\alpha_2\alpha_1\beta}(\omega_1,-\omega_{12})\right) = 0
\end{aligned} \tag{6.21}$$

Since Eq. 6.21 must be true for any $E(\omega_1)$ and $E(\omega_2)$,

$$(\omega_1+\omega_2)\,\chi_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) - \omega_1\,\chi_S^{\alpha_1\beta\alpha_2}(-\omega_1-\omega_2,\omega_2) - \omega_2\,\chi_S^{\alpha_2\alpha_1\beta}(\omega_1,-\omega_1-\omega_2) = 0 \tag{6.22}$$

which must also be respected for any frequencies ω_1 and ω_2, implying

$$\chi_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \chi_S^{\alpha_1\beta\alpha_2}(-\omega_1-\omega_2,\omega_2) = \chi_S^{\alpha_2\alpha_1\beta}(\omega_1,-\omega_1-\omega_2) \tag{6.23}$$

This is the condition of overall permutation symmetry. The subscript S serves as a reminder that only the part that respects intrinsic permutation symmetry has physical meaning and must therefore obey Eq. 6.23.

Before transposing this result to the language of conductivities, it is worth noting that overall permutation symmetry is a generalization of intrinsic permutation symmetry. To evidence this, an often employed notation for the susceptibility, and the conductivity, includes an additional argument to the function that is minus the sum of the other frequency arguments.

$$\chi^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \to \chi^{\beta\alpha_1\alpha_2}(-\omega_{12};\omega_1,\omega_2) \tag{6.24}$$

This notation is more natural than it may seem, since $E^{\beta}(\omega_{12})$ is the emitted optical field as a result of energy absorption from the incident fields $E^{\alpha_1}(\omega_1)$ and $E^{\alpha_2}(\omega_2)$. Making use of this notation in Eq. 6.23,

$$\chi_S^{\beta\alpha_1\alpha_2}(-\omega_{12};\omega_1,\omega_2) = \chi_S^{\alpha_1\beta\alpha_2}(\omega_1;-\omega_{12},\omega_2) = \chi_S^{\alpha_2\alpha_1\beta}(\omega_2;\omega_1,-\omega_{12}) \tag{6.25}$$

Overall permutation symmetry states that the susceptibility must be invariant under the permutation of not only the index pairs $\{\alpha_1,\omega_1\}$ and $\{\alpha_2,\omega_2\}$, as dictated by intrinsic permutation symmetry, but also with $\{\beta,-\omega_{12}\}$.

[4]For instance, the last term in Eq. 6.20 is the result of exchanging the dummy indices β and α_2, together with changing the integration variable $\omega_2 \to -\omega_1-\omega_2$.

The linear order result (Eq. 6.18) can be revisited in light of this new perspective,

$$\chi^{\beta\alpha}(-\omega;\omega) = \chi^{\alpha\beta}(\omega;-\omega) \tag{6.26}$$

and the condition of hermiticity recognized as a combination of overall permutation symmetry, which consists of the free exchange of the index pairs $\{\beta,-\omega\}$ and $\{\alpha,\omega\}$, and the reality condition.

Despite being useful for the purpose of transparently expressing permutation symmetries, the $\chi^{\beta\alpha_1\alpha_2}(-\omega_{12};\omega_1,\omega_2)$ notation introduces an redundant input to the response function and it was opted in this thesis to adopted the more succinct $\chi^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)$ notation.

Eq. 6.23 can be stated in terms of conductivities,

$$\frac{\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)}{\omega_1+\omega_2} = -\frac{\sigma_S^{\alpha_1\beta\alpha_2}(-\omega_1-\omega_2,\omega_2)}{\omega_1} = -\frac{\sigma_S^{\alpha_2\alpha_1\beta}(\omega_1,-\omega_1-\omega_2)}{\omega_2} \tag{6.27}$$

Eq. 6.27 might seem like a strange identity for the optical conductivity to obey, but a direct physical interpretation can be made of overall permutation symmetry if we look a bit past our semiclassical treatment. In a nonlinear medium with incident light waves $E^{\alpha_1}(\omega_1)$ and $E^{\alpha_2}(\omega_2)$, the energy absorbed, per unit volume, per unit time, at a frequency $\omega = \omega_1+\omega_2$ is given by $\sigma^{\beta\alpha_1\alpha_2}(\omega_1+\omega_2)(E^{\beta}(\omega_1+\omega_2))^* E^{\alpha_1}(\omega_1)E^{\alpha_2}(\omega_2)$. The energy of a single photon at the same frequency is $\hbar\omega$. The ratio of these two energies is the density of photons absorbed at ω and is what appears in the overall permutation symmetry condition. Eq. 6.27 can therefore be interpreted as stating that the number of photons emitted at $\omega_1+\omega_2$ is the same as the number of photons absorbed at ω_1 (or at ω_2), as it is to be expected from a simple energy level diagram (Fig. 1.1b). Closely related to this interpretation are the Manley-Rowe relations, that provide a more rigorous statement of energy conservation in a lossless medium and can be derived from overall permutation symmetry [27].

In solids, the condition for the medium to be lossless is that the incoming photons must not have sufficient energy to excite electrons from the valence bands to the conduction bands. Overall permutation symmetry, as it is commonly stated (Eqs. 6.17 and 6.27), is then observed only if the photon frequencies lie below the band gap Δ: $|\omega|<\Delta$ in linear order and $|\omega_1|<\Delta \wedge |\omega_2|<\Delta \wedge |\omega_1|+|\omega_2|<\Delta$ in second order. Additionally, no free charge carriers must be present: the Fermi level is set somewhere in the gap[5].

As a side note, the Kleinman symmetry alluded to in Section 1.2 is a special case of overall permutation symmetry, where the considered optical frequencies lie so far below the band gap that the dispersion is negligible. Then,

$$\chi_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \simeq \chi_S^{\alpha_1\beta\alpha_2}(\omega_1,\omega_2) \simeq \chi_S^{\alpha_1\alpha_2\beta}(\omega_1,\omega_2) \tag{6.28}$$

significantly reducing the number of independent tensor components. This condition must always be true for insulators in the DC limit $\omega_1,\omega_2 \to 0$.

It is possible to extend overall permutation symmetry to metals and doped semiconductors, which possess a Fermi surface, with minor adjustments. Instead of the band gap, it matters to consider the effective gap set by the chemical potential.

[5] In the language of the previous chapter, $\sigma_F = 0$.

Also, resonances at zero frequency must be avoided: DC conductivities are not expected to respect the symmetry, adding the conditions $\omega \neq 0$, at linear order, and $\omega_1 \neq 0 \neq \omega_2 \wedge \omega_1 + \omega_2 \neq 0$, at second order. Kleinman symmetry can be considered valid in the DC limit, as long as the optical frequencies remain sufficiently larger that the relaxation parameter γ.

Overall permutation symmetry can also be derived for third order, where it takes the form[6]

$$\begin{aligned} \frac{\sigma_S^{\beta\alpha_1\alpha_2\alpha_3}(\omega_1,\omega_2,\omega_3)}{\omega_{123}} &= -\frac{\sigma_S^{\alpha_1\beta\alpha_2\alpha_3}(-\omega_{123},\omega_2,\omega_3)}{\omega_1} \\ = -\frac{\sigma_S^{\alpha_2\alpha_1\beta\alpha_3}(\omega_1,-\omega_{123},\omega_3)}{\omega_2} &= -\frac{\sigma_S^{\alpha_3\alpha_1\alpha_2\beta}(\omega_1,\omega_2,-\omega_{123})}{\omega_3} \end{aligned} \tag{6.29}$$

for frequencies which, both individually and summed, are smaller than the gap[7]. The generalization to higher orders is obvious.

Notice that, both in second and third orders, only the first equality in our statements of overall permutation symmetry (Eqs. 6.27 and 6.29) is strictly necessary, the rest following from intrinsic permutation symmetry.

The most remarkable aspect of overall permutation symmetry is that it connects different frequency components of the nonlinear conductivity. As such, it provides non-trivial connections between distinct nonlinear optical effects (Section 1.1). So far, the validity of these relations is limited to the regime of sufficient low frequencies, in order to ensure there are no optical losses. It is also reliant on the absence of relaxation. In what follows, overall permutation symmetry will be generalized and these assumptions dropped.

6.3 Overall permutation symmetry

In the previous section, overall permutation symmetry was seen as a result of having zero optical losses in the system. It is, however, possible to rephrase the symmetry in a way that extends its generality.

The extension of overall permutation symmetry to lossy media has conceptual subtleties that we shall soon discuss, but is trivial to state: by analytical continuation of Eqs. 6.17, 6.27 and 6.29 into the Argand plane (Fig. 6.1), they become valid for any medium. Simply replacing $\omega \to \bar{\omega}$,

$$\sigma^{\beta\alpha}(\bar{\omega}) = -\sigma^{\alpha\beta}(-\bar{\omega}) \tag{6.30}$$

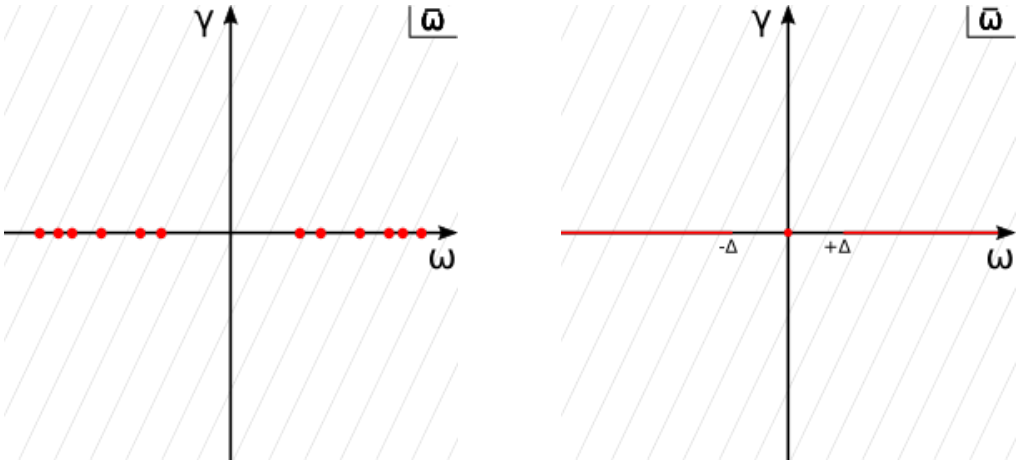
$$\frac{\sigma_S^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)}{\bar{\omega}_{12}} = -\frac{\sigma_S^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2)}{\bar{\omega}_1} \tag{6.31}$$

$$\frac{\sigma_S^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3)}{\bar{\omega}_{123}} = -\frac{\sigma_S^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_{123},\bar{\omega}_2,\bar{\omega}_3)}{\bar{\omega}_1} \tag{6.32}$$

provides a generalized form of overall permutation symmetry, that works for any region in frequency space where the conductivity function is naturally defined, no

[6] Recall the definition: $\omega_{123} \equiv \omega_1 + \omega_2 + \omega_3$.

[7] $\forall\, i,j,k \in \{1,2,3\}, |\omega_i| < \Delta \wedge |\omega_j| + |\omega_k| < \Delta \wedge |\omega_1| + |\omega_2| + |\omega_3| < \Delta$.



(a) (b)

Figure 6.1: The figures above depict the domain of the optical conductivity, with $\omega = \mathrm{Re}\{\bar{\omega}\}$ and $\gamma = \mathrm{Im}\{\bar{\omega}\}$. The resonances are outlined in red for the two cases: (a) an atomic system, where resonances occur when the optical frequency matches discrete energy levels; (b) a crystal, where resonances are available at any photon frequency that exceeds the gap Δ_{eff}. In a realistic band structure, other gaps may open up at higher frequencies. The resonance at the origin corresponds to the Drude peak, caused by the intraband motion of free charge carriers, and disappears in the absence of a Fermi surface. In both (a) and (b), there are no optical losses for any real frequency that is not marked in red.

matter if the optical frequencies lie below or above the gap, if electron transitions cause any optical losses or not.

Overall permutation symmetry was first expressed and analyzed in this form by Butcher and McLean [27]. Hereafter, when mentioning overall permutation symmetry, it will be in reference to the most general form of this symmetry, as stated in Eqs. 6.30-6.32.

To understand the fundamental difference on how this symmetry applies to lossless and lossy media, it is helpful to inspect the domain of the conductivity function. The easiest case to visualize is, of course, the complex plane $\mathbb{C}$ in linear order, depicted in Fig. 6.1. In red, the regions where resonances occur are outlined. These are the regions where the expressions for the optical conductivity are strictly undefined[8]. For crystals, resonances occur all along the real line, as soon as the photon frequency exceeds the band gap: $|\omega| > \Delta$. If the system has free carriers, the effective gap takes the role of the band gap and an additional resonance appears at the origin, due to intraband motion.

The segment of the real line where the optical conductivity is defined is the region of the Argand plane assigned to lossless media. In the previous section, it was demonstrated by an energy conservation argument that overall permutation symmetry was obeyed by the conductivity precisely in this region. Analytical continuation suggests the validity of the relations in Section 6.2 extends to the rest of the complex plane, with the exception of the cut along the real line (and potentially the origin).

It had already been established that complex frequencies were a requirement

[8]This applies to any expression derived in this thesis, be it by length or velocity gauge, at any level of generality (Chapters 2, 4 and 5).

when adopting optical frequencies above the gap (Section 2.5). It is then only natural to consider this extension of overall permutation symmetry. It had, however, also been pointed out that the upper and lower halves of the complex plane do not have quite the same meaning. The conductivity that has a physical interpretation, as the frequency domain response function derived from Eq. 2.77, inhabits the upper half of the complex plane. For this reason, it was mentioned that it was actually possible to define the optical conductivity for real frequencies above the gap, by considering the limit from above,

$$\sigma^{\beta\alpha}(\omega) \equiv \lim_{\gamma\to 0^+} \sigma^{\beta\alpha}(\bar{\omega}) \tag{6.33}$$

The difficulty, patent in Eqs. 6.30-6.32, lies in the fact that overall permutation symmetry connects the two halves of the complex plane, the "physical" and the "unphysical" conductivity. Only when their values agree on approaching the real line is overall permutation symmetry valid for real frequencies. In general, this is not true.

$$\lim_{\gamma\to 0^+} \sigma^{\beta\alpha}(\bar{\omega}) \neq \lim_{\gamma\to 0^-} \sigma^{\beta\alpha}(\bar{\omega}) \qquad \text{for } |\omega| > \Delta \tag{6.34}$$

This will be made explicit in the next section. Here, it is enough to know that the two limits match only below the gap, precisely for the case of a lossless system. Once the photon frequency exceeds the gap, causing real electronic transitions in the system, the conductivity in the limit of vanishing relaxation, defined by Eq. 6.33, no longer respects the relations of Section 6.2,

$$\sigma^{\beta\alpha}(\omega) \neq -\sigma^{\alpha\beta}(-\omega) \qquad \text{for } |\omega| > \Delta \tag{6.35}$$

A similar reasoning applies to the nonlinear conductivity (Section 6.4).

The validity of the statements of overall permutation symmetry (Eqs. 6.30-6.32), from linear to third order, can be verified by direct substitution, using the expressions obtained from the length gauge perturbation theory for the nonlinear conductivity[9], as presented in Section 5.2.

As an example, this is shown explicitly for the linear conductivity:

$$\begin{aligned}
-\sigma^{\alpha\beta}(-\bar{\omega}) &= -\frac{i\,e^2}{\hbar}\left(\frac{1}{\hbar\bar{\omega}}\,F_A^{\alpha\beta} + i\,F_B^{\alpha\beta} - \hbar\bar{\omega}\,\Pi_1^{\alpha\beta}(-\bar{\omega})\right) \\
&= \frac{i\,e^2}{\hbar}\left(-\frac{1}{\hbar\bar{\omega}}\,F_A^{\alpha\beta} - i\,F_B^{\alpha\beta} + \hbar\bar{\omega}\,\Pi_1^{\alpha\beta}(-\bar{\omega})\right) \\
&= \frac{i\,e^2}{\hbar}\left(-\frac{1}{\hbar\bar{\omega}}\,F_A^{\beta\alpha} + i\,F_B^{\beta\alpha} + \hbar\bar{\omega}\,\Pi_1^{\beta\alpha}(\bar{\omega})\right) \\
&= \sigma^{\beta\alpha}(\bar{\omega})
\end{aligned} \tag{6.36}$$

where integral identities were used (Eqs. G.17 and G.18), including the statement $\Pi_1^{\beta\alpha}(-\bar{\omega}) = \Pi_1^{\alpha\beta}(\bar{\omega})$, easily proved from the definition of the integral (Eq. 5.21).

[9] Checking that the conductivities in Section 5.2 respect overall permutation symmetry is not as straightforward as it may seem, due to identities that result from the commutativity of covariant derivatives (see Appendix C; the relation between integrals in Eq. C.9 is particularly troublesome). On the other hand, overall permutation symmetry is trivially checked in the expressions in Appendix E, where the commutativity of covariant derivatives was abandoned.

Likewise, the overall permutation symmetry of the second and third order conductivities can be attested, with the use of Eqs. 5.22-5.42 and the respective integral identities (Eqs. G.20-G.31 in Appendix G). It requires writing more extensive equations to check overall permutation symmetry at second and third orders, if for nothing else, because σ_S must used, with prior symmetrization to ensure intrinsic permutation symmetry.

Even though such direct proof uses the equations from Section 5.2, derived for the two-band crystal, the nature of these relations and their connection to the fundamental principle of energy conservation (Section 6.2) invites the conjecture that they must be valid for any number of bands and even beyond third order perturbation theory. A general proof is given in [27] using the velocity gauge (in its older form, as discussed in Section 2.4.2), that also indicates that this symmetry is not constrained to the examples here demonstrated. Overall permutation symmetry can be expected to hold true for any nonlinear conductivity.

This is a rather interesting statement, considering that, even after careful inspection of Eq. 2.64 (or Eq. 4.25), it could hardly be guessed that this expression, after proper symmetrization, would respect Eq. 6.32. From this perspective, these relations constitute a very non-obvious property of the nonlinear conductivity.

Also, the overall permutation symmetry obeyed by the nonlinear susceptibility is simpler, perhaps more natural, than the one presented by the nonlinear conductivity. Why is this the case? In fact, it is a slight misuse of language to state, as is done in this thesis, that the conductivity respects overall permutation symmetry, since this designation is more properly applied to Eqs. 6.18 and 6.23, of which the relations for the conductivity are a consequence.

If we return to the perspective of Sections 2.1 and 2.2, that developed the perturbative treatment for a general response function, relating an observable of interest J to a classical field that coupled to another system observable $\mathcal{O}$, we recognize that there is something very special in the case of the susceptibility. For the susceptibility, the observable of interest is the polarization $-e\,\hat{r}$ induced by the coupling of the optical field to the dipole operator $\hat{\mathcal{O}} = e\,\hat{r}$. In other words, when developing the perturbation theory for the susceptibility[10], in the terminology of Section 2.2, we have $J = -\mathcal{O}$: *aside from an irrelevant minus sign, the observable of interest is the one that couples to the classical field.*

It follows that whenever this condition is met and the interest of a physicist falls upon an observable that couples to a classical field, weakly enough for a perturbative approach to be applicable, the response functions that determine this observable will respect overall permutation symmetry. The same conclusion could be derived by energy considerations, akin to Section 6.2. It is then a wider symmetry, well-known within the confines of nonlinear optics, but that survives in other contexts.

In the case of nonlinear optics, Eqs. 6.30-6.32 establish relationships between distinct nonlinear optical effects. The most interesting cases are listed in Table 6.1. The connected effects may require experimental apparatus operating in distant re-

[10]Set aside the fact that this is not quite so simple to do for nonlinear optics in solids, due to the position operator taking the form of a covariant derivative. While in the perturbative treatment of the electric current, the position operator appears always inside commutators, in the perturbative expansion of the polarization this is not the case and, in actuality, it is not even possible to a adopt a single-particle picture: second quantization notation is required to write the polarization operator in the Bloch basis [5]. See Appendix B for more on the connection between the susceptibility and the conductivity.

$\sigma(\omega,\omega)$ Second harmonic generation	$\sigma(-2\omega,\omega)$ Difference frequency generation
$\sigma(\omega,0)$ Pockels effect	$\sigma(\omega,-\omega)$ Shift and injection currents
$\sigma(\omega,\omega,\omega)$ Third harmonic generation	$\sigma(-3\omega,\omega,\omega)$ Difference frequency generation
$\sigma(\omega,\omega,0)$ DC field induced second harmonic generation	$\sigma(\omega,\omega,-2\omega)$ Two color current injection
$\sigma(\omega,0,0)$ Quadratic electro-optic effect	$\sigma(\omega,-\omega,0)$ Jerk, injection and shift currents
$\sigma(\omega,-\omega,\Omega)$ Optical Kerr effect: light wave at ω modulates refractive index at Ω	$\sigma(\Omega,-\Omega,\omega)$ Optical Kerr effect: light wave at Ω modulates refractive index at ω

Table 6.1: Overall permutation symmetry provides surprising links between nonlinear optical effects. The conductivities on the right can be derived from the ones presented on the left by permutation symmetry and vice versa. Tensor indices were omitted.

gions of the electromagnetic spectrum, since they concern different components of the electric current, and are often of interest to different research communities. The way they are intimately related by this symmetry demonstrates the advantage of having an unified treatment, provided by the nonlinear conductivity.

6.4 The dissipative part

The most relevant situation, physically, is when the phenomenological parameter γ is negligible. This is when the features in the dispersion of the nonlinear conductivity are more pronounced and when analytical derivations become possible (see Section 5.3). It is also when the "phenomenological" part becomes less important.

In the previous section, the relaxation-free limit of the relations set by overall permutation symmetry was addressed in somewhat vague terms and it was argued that these relations would be violated in this limit, above the effective gap, due to a mismatch in the limits that approach real frequencies from above and below in the complex plane. This question is examined in greater detail here, by pulling on the

results of Chapter 5.

For simplicity, we consider a gapped two-band system and assume that the Fermi level resides in the gap ($\sigma_F = 0$). In which case (Eqs. 5.16 and 5.18),

$$\sigma^{\beta\alpha}(\bar{\omega}) = i\, e^2\, \bar{\omega}\, \Pi_1^{\beta\alpha}(\bar{\omega}) \tag{6.37}$$

In the limit of vanishing relaxation, using Eq. 5.44,

$$\sigma^{\beta\alpha}(\omega) \equiv \lim_{\gamma\to 0^+} \sigma^{\beta\alpha}(\bar{\omega}) = \pi\, i\, e^2\, \omega \left(\mathcal{H}_1^{\beta\alpha}(\omega) - i\,\mathcal{I}_1^{\beta\alpha}(\omega)\right) = \sigma_{\mathcal{H}}^{\beta\alpha}(\omega) + \sigma_{\mathcal{I}}^{\beta\alpha}(\omega) \tag{6.38}$$

where the notation $\sigma_{\mathcal{I}}$ and $\sigma_{\mathcal{H}}$ is introduced, representing the portions of the conductivity given by the $\mathcal{I}_j$ integrals and their Hilbert transforms $\mathcal{H}_j$, respectively.

Overall permutation symmetry relates Eq. 6.37 to another frequency component, $\sigma(-\bar{\omega})$, for which the limit gives a different result,

$$\lim_{\gamma\to 0^+} \sigma^{\beta\alpha}(-\bar{\omega}) = \pi\, i\, e^2\, \omega \left(\mathcal{H}_1^{\beta\alpha}(-\omega) + i\,\mathcal{I}_1^{\beta\alpha}(-\omega)\right) = \sigma_{\mathcal{H}}^{\beta\alpha}(-\omega) - \sigma_{\mathcal{I}}^{\beta\alpha}(-\omega) \tag{6.39}$$

This can be understood by returning to Fig. 6.1 and recognizing that, due to the sign of the imaginary part being reversed $\sigma(-\bar{\omega}) = \sigma(-\omega - i\gamma)$, the real frequencies are approached from below in the complex plane, instead of above. As a result, when the Sokhotski–Plemelj theorem is employed in expanding the denominator in Eq. 5.6, a sign flip will appear in the $\mathcal{I}_1$ term, as already described in Eq. 5.47.

In order to avoid confusion, it is important to remember that the physical conductivity $\sigma(\omega)$ is defined by taking the limit from above and that

$$\sigma^{\beta\alpha}(-\omega) \neq \lim_{\gamma\to 0^+} \sigma^{\beta\alpha}(-\bar{\omega}) \tag{6.40}$$

With Eqs. 6.38 and 6.39, we arrive at the limit $\gamma \to 0^+$ of Eq. 6.30,

$$\sigma_{\mathcal{H}}^{\beta\alpha}(\omega) = -\sigma_{\mathcal{H}}^{\alpha\beta}(-\omega) \tag{6.41}$$

$$\sigma_{\mathcal{I}}^{\beta\alpha}(\omega) = \sigma_{\mathcal{I}}^{\alpha\beta}(-\omega) \tag{6.42}$$

For light frequencies below the gap $|\omega| < \Delta$, the integral $\mathcal{I}_1$ is trivially zero and we have $\sigma_{\mathcal{I}} = 0$. This reduces to the already examined case of a lossless medium, since $\sigma = \sigma_{\mathcal{H}}$ is anti-hermitian (Eq. 6.17).

However, if we have photons of a higher energy, $|\omega| > \Delta$, then resonances will take place, the $\mathcal{I}$ integrals are no longer negligible and the $\sigma_{\mathcal{I}}$ portion of the conductivity will break the anti-hermiticity:

$$-\sigma^{\alpha\beta}(-\omega) = -\sigma_{\mathcal{H}}^{\alpha\beta}(-\omega) - \sigma_{\mathcal{I}}^{\alpha\beta}(-\omega) \neq \sigma_{\mathcal{H}}^{\beta\alpha}(\omega) + \sigma_{\mathcal{I}}^{\beta\alpha}(\omega) = \sigma^{\beta\alpha}(\omega) \tag{6.43}$$

Since, for real frequencies, overall permutation symmetry is synonymous with lossless propagation of light (Section 6.2), identifying the part of the conductivity that does not satisfy Eq. 6.17 corresponds to isolating the part that is responsible for optical absorption: $\sigma_{\mathcal{I}}$.

For the purposes of studying linear absorption, the computation of the $\mathcal{I}_1$ integral is sufficient. This is a standard technique: $\mathcal{I}_1$ corresponds to the traditional Fermi golden rule, which provides the transition rate and the absorption coefficient. Its Hilbert transform gives the refractive index, via the Kramers-Krönig relations (Section 6.1).

The same principles apply to the second order conductivity,

$$\sigma^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \equiv \lim_{\gamma\to 0^+} \sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) \tag{6.44}$$

The relation in Eq. 6.27, derived from conservation of energy in a lossless nonlinear medium, does not hold for optical frequencies above the gap. Insight on this is gained by inspection of the relaxation free-limit of Eq. 6.31. The relation concerns the symmetrized second order conductivity,

$$\begin{aligned}\sigma_S^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) &= \frac{1}{2}\left(\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)\right)\\ &= \frac{1}{2}\left(\sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma_1^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)+\sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma_2^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)\right)\end{aligned} \tag{6.45}$$

Eq. 5.47 is used to evaluate the relaxation-free limit of the one-photon contribution (Eq. 5.24),

$$\sigma_1^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \lim_{\gamma\to 0^+} \sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_{1\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)+\sigma_{1\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \tag{6.46}$$

where

$$\frac{\hbar}{\pi\, i\, e^3}\,\sigma_{1\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{1}{\hbar\omega_1+\hbar\omega_2}\,\mathcal{H}_1^{\alpha_2\alpha_1\beta}(\omega_1)+\frac{\hbar\omega_1+\hbar\omega_2}{(\hbar\omega_2)^2}\,\mathcal{H}_1^{\beta\alpha_1\alpha_2}(\omega_1)+\frac{\hbar\omega_1}{\hbar\omega_2}\,\mathcal{H}_2^{\alpha_1\beta\alpha_2}(-\omega_1) \tag{6.47}$$

$$\frac{\hbar}{\pi\, e^3}\,\sigma_{1\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{1}{\hbar\omega_1+\hbar\omega_2}\,\mathcal{I}_1^{\alpha_2\alpha_1\beta}(\omega_1)+\frac{\hbar\omega_1+\hbar\omega_2}{(\hbar\omega_2)^2}\,\mathcal{I}_1^{\beta\alpha_1\alpha_2}(\omega_1)-\frac{\hbar\omega_1}{\hbar\omega_2}\,\mathcal{I}_2^{\alpha_1\beta\alpha_2}(-\omega_1) \tag{6.48}$$

and the two-photon contribution (Eq. 5.25),

$$\sigma_2^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \lim_{\gamma\to 0^+} \sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_{2\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)+\sigma_{2\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \tag{6.49}$$

where

$$\frac{\hbar}{\pi\, i\, e^3}\,\sigma_{2\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = -\frac{\hbar\omega_1+\hbar\omega_2}{(\hbar\omega_2)^2}\,\mathcal{H}_1^{\beta\alpha_1\alpha_2}(\omega_1+\omega_2)+\frac{\hbar\omega_1+\hbar\omega_2}{\hbar\omega_2}\,\mathcal{H}_2^{\beta\alpha_1\alpha_2}(\omega_1+\omega_2) \tag{6.50}$$

$$\frac{\hbar}{\pi\, e^3}\,\sigma_{2\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = -\frac{\hbar\omega_1+\hbar\omega_2}{(\hbar\omega_2)^2}\,\mathcal{I}_1^{\beta\alpha_1\alpha_2}(\omega_1+\omega_2)+\frac{\hbar\omega_1+\hbar\omega_2}{\hbar\omega_2}\,\mathcal{I}_2^{\beta\alpha_1\alpha_2}(\omega_1+\omega_2) \tag{6.51}$$

In the relaxation-free limit, by which it is always meant $\gamma \to 0^+$, the nonlinear conductivity may diverge for specific frequency components (this does happen for most of the effects in Table 6.1), but it was assumed in the reasoning above that this is not the case. As discussed in Section 5.3, the direct application of Eq. 5.47 to evaluate the relaxation-free limit, without concern for how the coefficients C multiplying the Π_j integrals behave in this limit, presupposes that we exclude certain regions of frequency space from our analysis. In this case, we must assume that $\omega_1 \neq 0 \neq \omega_2$ and $\omega_1 + \omega_2 \neq 0$. These situations are of physical interest, but, as mentioned before, the topic of the singularities of the nonlinear conductivity resides outside the scope of this thesis.

Having assessed the limits for the one-photon and two-photon contributions, we can now consider the entire second order conductivity,

$$\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \equiv \lim_{\gamma\to 0^+} \sigma_S^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_{\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma_{\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \tag{6.52}$$

where we summed up all contributions involving the $\mathcal{H}$ and $\mathcal{I}$ integrals into

$$\sigma_{\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{1}{2}\left(\sigma_{1\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma_{1\mathcal{H}}^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1) + \sigma_{2\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma_{2\mathcal{H}}^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1)\right) \tag{6.53}$$

$$\sigma_{\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \frac{1}{2}\left(\sigma_{1\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma_{1\mathcal{I}}^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1) + \sigma_{2\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) + \sigma_{2\mathcal{I}}^{\beta\alpha_2\alpha_1}(\omega_2,\omega_1)\right) \tag{6.54}$$

respectively.

Overall permutation symmetry (Eq. 6.31) connects this to another frequency component, $\sigma_S^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2)$, whose resonance-based decomposition gives

$$\begin{aligned}\sigma_S^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) &= \frac{1}{2}\left(\sigma^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) + \sigma^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12})\right) \\ &= \frac{1}{2}\left(\sigma_1^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) + \sigma_1^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12}) + \sigma_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) + \sigma_2^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12})\right)\end{aligned} \tag{6.55}$$

Similarly to before, we first obtain the relaxation free limit of the one-photon contribution,

$$\lim_{\gamma\to 0^+} \sigma_1^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12}) = \sigma_{1\mathcal{H}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) + \sigma_{1\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) \tag{6.56}$$

The issue is in the other one-photon contribution, related to the photon frequency $-\bar{\omega}_{12}$. The complex frequency argument in the Π_j integrals will have a minus sign, therefore

$$\lim_{\gamma\to 0^+} \sigma_1^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) = \sigma_{1\mathcal{H}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) - \sigma_{1\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) \tag{6.57}$$

As it happened in linear order, there a sign flip in the contribution involving the $\mathcal{I}_j$ integrals. The same will occur for the remaining two-photon contributions, associated with the frequency $-\bar{\omega}_1 = -\bar{\omega}_{12} + \bar{\omega}_2$,

$$\lim_{\gamma\to 0^+} \sigma_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) = \sigma_{2\mathcal{H}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) - \sigma_{2\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) \tag{6.58}$$

$$\lim_{\gamma\to 0^+} \sigma_2^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12}) = \sigma_{2\mathcal{H}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) - \sigma_{2\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) \tag{6.59}$$

Combining the previous limits,

$$\begin{aligned}
&\lim_{\gamma\to 0^+} \sigma_S^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) \\
&= \lim_{\gamma\to 0^+} \frac{1}{2}\left(\sigma_1^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) + \sigma_1^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12}) + \sigma_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) + \sigma_2^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12})\right) \\
&= + \frac{1}{2}\left(\sigma_{1\mathcal{H}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) + \sigma_{1\mathcal{H}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) + \sigma_{2\mathcal{H}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) + \sigma_{2\mathcal{H}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12})\right) \\
&\quad + \frac{1}{2}\left(-\sigma_{1\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) + \sigma_{1\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) - \sigma_{2\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) - \sigma_{2\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12})\right)
\end{aligned} \tag{6.60}$$

from which we conclude,

$$\sigma_S^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) \neq \lim_{\gamma\to 0^+} \sigma_S^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2) \tag{6.61}$$

This is the second order analogue of Eq. 6.40. Again we find that, because overall permutation symmetry involves frequency components that approach the real frequencies from below in the complex plane, and the physical conductivity is defined from above, the relaxation-free limit of Eq. 6.31 is evaluated differently in the left and on the right hand sides, producing a mismatch in the signs of the portion of the conductivity that involves the $\mathcal{I}_j$ integrals.

Joining Eqs. 6.52 and 6.60, the relaxation-free limit of the overall permutation symmetry condition at second order is finally obtained,

$$\sigma_{\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = -\left(\frac{\omega_1+\omega_2}{\omega_1}\right)\sigma_{\mathcal{H}}^{\alpha_1\beta\alpha_2}(-\omega_1-\omega_2,\omega_2) \tag{6.62}$$

and

$$\begin{aligned}
\sigma_{\mathcal{I}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) &= -\left(\frac{\omega_{12}}{\omega_1}\right)\Big(-\sigma_{1\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) + \sigma_{1\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12}) \\
&\qquad - \sigma_{2\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) - \sigma_{2\mathcal{I}}^{\alpha_1\alpha_2\beta}(\omega_2,-\omega_{12})\Big) \\
&\neq -\left(\frac{\omega_{12}}{\omega_1}\right)\sigma_{\mathcal{I}}^{\alpha_1\beta\alpha_2}(-\omega_1-\omega_2,\omega_2)
\end{aligned} \tag{6.63}$$

The conclusions are identical to linear order. Below the band gap, the $\mathcal{I}_j$ integrals are trivially zero and only the Hilbert transforms survive. In this case, the second order conductivity $\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) = \sigma_{\mathcal{H}}^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2)$ obeys the relation in Eq. 6.27, that codifies the absence of any optical losses in the nonlinear medium.

Once any optical frequency exceeds the gap, the $\mathcal{I}_j$ integrals cannot be ignored and Eq. 6.27 is no longer valid,

$$\sigma_S^{\beta\alpha_1\alpha_2}(\omega_1,\omega_2) \neq -\left(\frac{\omega_{12}}{\omega_2}\right)\sigma_S^{\alpha_1\beta\alpha_2}(-\omega_{12},\omega_2) \tag{6.64}$$

This includes the possibilities of one-photon, $\omega_1 > \Delta \vee \omega_2 > \Delta$, or two-photon resonances, $\omega_1 + \omega_2 > \Delta$. Losses are introduced into the system by the $\sigma_{\mathcal{I}}^{\beta\alpha_1\alpha_2}$, identified as the dissipative part of the second order conductivity. In a generalization of the Fermi golden rule methodology, evaluation of the integrals $\mathcal{I}_1^{\beta\alpha_1\alpha_2}$ and $\mathcal{I}_2^{\beta\alpha_1\alpha_2}$ are all that is required to quantify the nonlinear absorption of the crystal.

We note that, curiously, $\sigma_{\mathcal{I}}$ does obey a very similar relation to overall permutation symmetry in linear order, with only a sign change (Eq. 6.42). But this seems to not be true for higher orders (see Eq. 6.63). Is it perhaps possible to derive some general symmetry that is respected by the dissipative part of the conductivity? To the author's knowledge, it is not. What is possible, though, is to construct an identity that is obeyed at second order, for diagonal matrix elements,

$$\sigma_{\mathcal{I}}^{\beta\beta\beta}(\omega_1,\omega_2) = \left(\frac{\omega_1+\omega_2}{\omega_1}\right)\sigma_{\mathcal{I}}^{\beta\beta\beta}(-\omega_1-\omega_2,\omega_2) - \left(\frac{\omega_2}{\omega_1}\right)\sigma_{\mathcal{I}}^{\beta\beta\beta}(\omega_2,-\omega_2) \tag{6.65}$$

This relation was first derived by Ventura and forms the second order analogue of Eq. 6.42. For more insight, and the actual derivation, see [96].

To conclude, we consider the changes caused by a Fermi level set above the gap. As usual, the effective gap replaces the band gap in previous statements, regarding resonances and optical loss. Other than that, the Fermi surface contributions σ_F satisfy Eqs. 6.17 and 6.27 for any nonzero frequency, implying they can just be added to the Hilbert transforms in Eqs. 6.41 and 6.62. Eqs. 6.17 and 6.27 are respected when all optical frequencies lie below the effective gap and all other conclusions are identical. There is the appearance of zero frequency resonances, but recall that these were excluded from our analysis anyway.

For brevity, this section covered the study of overall permutation symmetry in the relaxation-free limit up to second order, but the analysis is naturally extended to third order and Eqs. 6.29 and 6.32. The central ideas remain unchanged and can be applied to higher orders still. From the discussion above, it should be clear that the part of the conductivity containing the integrals $\mathcal{I}_j$ will be to blame for optical absorption, at any order in perturbation theory.

6.5 On a connection between photon resonances

The usefulness of the resonance-based analysis (Chapter 5) is not limited to the relaxation-free limit. To finish off this discussion on overall permutation symmetry, the resonance-based decomposition is performed on the overall permutation symmetry condition, from linear to third order, giving rise to a new set of identities. These provide unexpected connections among the various pieces that form the nonlinear conductivity.

As an warm up, we consider the linear case, for which the conclusions are not very surprising. The statement of overall permutation symmetry for the linear conductivity (Eq. 6.30) is combined with the decomposition in Eq. 5.16,

$$\sigma_F^{\beta\alpha}(\bar{\omega}) + \sigma_1^{\beta\alpha}(\bar{\omega}) = -\sigma_F^{\alpha\beta}(-\bar{\omega}) - \sigma_1^{\alpha\beta}(-\bar{\omega}) \tag{6.66}$$

The key step is to recognize that for Eq. 6.66 to be true for any crystal, the pole structure on the two sides must be matched. On each side, there is a term that is dictated by Fermi surface properties and another whose real part is fully determined by the electronic properties near the region of the FBZ where $\hbar\omega = \Delta\epsilon_{\mathbf{k}cv}$. Since these two regions of the FBZ are independent, and overall permutation symmetry must be satisfied for any band structure $\epsilon_{\mathbf{k}a}$ and Berry connection $\mathcal{A}^{\alpha}_{\mathbf{k}ab}$, Eq. 6.66 has to be verified piece-wise, with each contribution satisfying the anti-hermiticity condition separately,

$$\sigma_1^{\beta\alpha}(\bar{\omega}) = -\,\sigma_1^{\alpha\beta}(-\bar{\omega}) \tag{6.67}$$

$$\sigma_F^{\beta\alpha}(\bar{\omega}) = -\,\sigma_F^{\alpha\beta}(-\bar{\omega}) \tag{6.68}$$

Similarly, the Fermi surface contribution will be isolated in every order of perturbation theory. It is the only contribution that is related to the Fermi surface and not to a single or multi-photon resonance. In a sense, it could be said they involve zero-frequency resonances, but these are not of the type that is relevant here. The end result is that the Fermi surface contributions always obey overall permutation symmetry by themselves,

$$\sigma_{FS}^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = -\left(\frac{\bar{\omega}_1+\bar{\omega}_2}{\bar{\omega}_1}\right)\sigma_{FS}^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1-\bar{\omega}_2,\bar{\omega}_2) \tag{6.69}$$

$$\sigma_{FS}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = -\left(\frac{\bar{\omega}_1+\bar{\omega}_2+\bar{\omega}_3}{\bar{\omega}_1}\right)\sigma_{FS}^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_1-\bar{\omega}_2-\bar{\omega}_3,\bar{\omega}_2,\bar{\omega}_3) \tag{6.70}$$

where the 'S' subscript serves as a reminder that intrinsic permutation symmetry must be ensured.

The same cannot be said for the other contributions, when looking past linear order. In second order, introducing the resonance-based decomposition of Eq. 5.22 in Eq. 6.31,

$$\begin{aligned}&\frac{1}{2}\left(\sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma_1^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)\right)+\sigma_{2S}^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)\\&=-\left(\frac{\bar{\omega}_{12}}{2\,\bar{\omega}_1}\right)\left(\sigma_1^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2)+\sigma_1^{\alpha_1\alpha_2\beta}(\bar{\omega}_2,-\bar{\omega}_{12})+2\,\sigma_{2S}^{\alpha_1\beta\alpha_2}(-\bar{\omega}_{12},\bar{\omega}_2)\right)\end{aligned} \tag{6.71}$$

with

$$\sigma_{2S}^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) \equiv \frac{1}{2}\left(\sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2)+\sigma_2^{\beta\alpha_2\alpha_1}(\bar{\omega}_2,\bar{\omega}_1)\right) \tag{6.72}$$

The Fermi surface contributions were cancelled out from Eq. 6.71. Henceforth, they will be left out of the discussion, since their symmetry properties are already identified in Eqs. 6.69 and 6.70.

Recall that $\sigma_1(\bar{\omega}_1,\bar{\omega}_2)$ is ascribed to the one-photon resonance $\hbar\omega_1 = \Delta\epsilon_{\mathbf{k}cv}$ and $\sigma_2(\bar{\omega}_1,\bar{\omega}_2)$ to the two-photon resonance $\hbar\omega_1+\hbar\omega_2 = \Delta\epsilon_{\mathbf{k}cv}$ (assuming that ω_1,$\omega_2 > 0$). On the other hand, $\sigma_1(-\bar{\omega}_{12},\bar{\omega}_2)$ involves a resonance with a single photon of energy $\hbar\omega_{12}$, again giving $\hbar\omega_1+\hbar\omega_2 = \Delta\epsilon_{\mathbf{k}cv}$, while $\sigma_2(-\bar{\omega}_{12},\bar{\omega}_2)$ involves a two-photon resonance $-\hbar\omega_{12}+\hbar\omega_2 = -\hbar\omega_1 = \Delta\epsilon_{vc}$. The only way for Eq. 6.71 to hold is

for the contributions associated to specific regions of the FBZ (defined by a given resonance condition) to cancel separately. Recognizing this, Eq. 6.71 unfolds into a series of identities, that concern the resonances at $\hbar\omega_1 = \Delta\epsilon_{cv}$, $\hbar\omega_2 = \Delta\epsilon_{cv}$ and $\hbar\omega_1 + \hbar\omega_2 = \Delta\epsilon_{cv}$,

$$\sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1, \bar{\omega}_2) = -2\left(\frac{\bar{\omega}_1 + \bar{\omega}_2}{\bar{\omega}_1}\right)\sigma_{2S}^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1 - \bar{\omega}_2, \bar{\omega}_2) \tag{6.73}$$

$$\sigma_1^{\beta\alpha_2\alpha_1}(\bar{\omega}_2, \bar{\omega}_1) = -\left(\frac{\bar{\omega}_1 + \bar{\omega}_2}{\bar{\omega}_1}\right)\sigma_1^{\alpha_1\alpha_2\beta}(\bar{\omega}_2, -\bar{\omega}_1 - \bar{\omega}_2) \tag{6.74}$$

$$\sigma_{2S}^{\beta\alpha_1\alpha_2}(\bar{\omega}_1, \bar{\omega}_2) = -\left(\frac{\bar{\omega}_1 + \bar{\omega}_2}{2\,\bar{\omega}_1}\right)\sigma_1^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1 - \bar{\omega}_2, \bar{\omega}_2) \tag{6.75}$$

respectively.

Eq. 6.74 is interesting in how it constrains the space of functions that could serve as the one-photon contribution in a nonlinear conductivity. It is a symmetry of this type of contribution, at second order.

But the really interesting result is presented in the other equations. Eq. 6.73 and Eq. 6.75 are equivalent, as can be seen by a simple variable transformation, and tell us that *the one- and two-photon contributions to the second order conductivity are intimately related*. In fact, σ_1 can be obtained from σ_2 and vice-versa!

The implication is that, once the behavior of the second order conductivity at the two-photon resonance is known, the rest can be reconstructed on the basis of overall permutation symmetry. It follows that only one of the two pieces in the resonance-based decomposition of the second order conductivity needs to be computed.

From the practical standpoint, based on the author's experience, Eqs. 6.73-6.75 serve mainly as a helpful consistency check in the elaborate computations necessary to obtain a crystal's nonlinear conductivity.

A deep connection between different resonances can be found in the third order nonlinear conductivity as well. The symmetrized form of the third order conductivity, written in detail, is

$$\begin{aligned}
&\sigma_S^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) \\
\equiv &+\frac{1}{6}\Big(\sigma_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma_1^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}_1, \bar{\omega}_3, \bar{\omega}_2) + \sigma_1^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2, \bar{\omega}_3, \bar{\omega}_1) \\
&\qquad + \sigma_1^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}_2, \bar{\omega}_1, \bar{\omega}_3) + \sigma_1^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3, \bar{\omega}_1, \bar{\omega}_2) + \sigma_1^{\beta\alpha_3\alpha_2\alpha_1}(\bar{\omega}_3, \bar{\omega}_2, \bar{\omega}_1)\Big) \\
&+\frac{1}{6}\Big(\sigma_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma_2^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}_1, \bar{\omega}_3, \bar{\omega}_2) + \sigma_2^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2, \bar{\omega}_3, \bar{\omega}_1) \\
&\qquad + \sigma_2^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}_2, \bar{\omega}_1, \bar{\omega}_3) + \sigma_2^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3, \bar{\omega}_1, \bar{\omega}_2) + \sigma_2^{\beta\alpha_3\alpha_2\alpha_1}(\bar{\omega}_3, \bar{\omega}_2, \bar{\omega}_1)\Big) \\
&+\frac{1}{6}\Big(\sigma_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma_3^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}_1, \bar{\omega}_3, \bar{\omega}_2) + \sigma_3^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2, \bar{\omega}_3, \bar{\omega}_1) \\
&\qquad + \sigma_3^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}_2, \bar{\omega}_1, \bar{\omega}_3) + \sigma_3^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3, \bar{\omega}_1, \bar{\omega}_2) + \sigma_3^{\beta\alpha_3\alpha_2\alpha_1}(\bar{\omega}_3, \bar{\omega}_2, \bar{\omega}_1)\Big) \\
= &+\frac{1}{3}\Big(\sigma_{1S}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma_{1S}^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2, \bar{\omega}_3, \bar{\omega}_1) + \sigma_{1S}^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3, \bar{\omega}_1, \bar{\omega}_2)\Big) \\
&+\frac{1}{3}\Big(\sigma_{2S}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) + \sigma_{2S}^{\beta\alpha_2\alpha_3\alpha_1}(\bar{\omega}_2, \bar{\omega}_3, \bar{\omega}_1) + \sigma_{2S}^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_3, \bar{\omega}_1, \bar{\omega}_2)\Big) \\
&+\sigma_{3S}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3)
\end{aligned} \tag{6.76}$$

where $\sigma_{1S}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3)$ is symmetrized with respect to the last two index pairs, $\{\alpha_2, \bar{\omega}_2\}$ and $\{\alpha_3, \bar{\omega}_3\}$, $\sigma_{2S}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3)$ is symmetrized with respect to the first two index pairs and $\sigma_{3S}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3)$ is symmetrized with respect to all index pairs.

The overall permutation symmetry of Eq. 6.32, after performing the resonance-based decomposition in Eq. 6.76, and by the same arguments as before, unfolds into a series of fundamental identities,

$$\sigma_{1S}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = -3\left(\frac{\bar{\omega}_1 + \bar{\omega}_2 + \bar{\omega}_3}{\bar{\omega}_1}\right) \sigma_{3S}^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_1 - \bar{\omega}_2 - \bar{\omega}_3, \bar{\omega}_2, \bar{\omega}_3) \quad (6.77)$$

$$\sigma_{1S}^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}_2, \bar{\omega}_1, \bar{\omega}_3) = -\left(\frac{\bar{\omega}_1 + \bar{\omega}_2 + \bar{\omega}_3}{\bar{\omega}_1}\right) \sigma_{1S}^{\alpha_1\alpha_2\beta\alpha_3}(\bar{\omega}_2, -\bar{\omega}_1 - \bar{\omega}_2 - \bar{\omega}_3, \bar{\omega}_3) \quad (6.78)$$

$$\sigma_{2S}^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1, \bar{\omega}_2, \bar{\omega}_3) = -\left(\frac{\bar{\omega}_1 + \bar{\omega}_2 + \bar{\omega}_3}{\bar{\omega}_1}\right) \sigma_{2S}^{\alpha_1\beta\alpha_3\alpha_2}(-\bar{\omega}_1 - \bar{\omega}_2 - \bar{\omega}_3, \bar{\omega}_3, \bar{\omega}_2) \quad (6.79)$$

together with other equivalent equations.

These relations are the expression of overall permutation symmetry for the separate contributions to the third order conductivity (as identified in Eq. 5.30). We find that Eqs. 6.78 and 6.79 place constraints on the shape of one-photon and two-photon contributions, σ_{1S} and σ_{2S}, respectively. This time, the link between resonances occurs between one-photon and three-photon terms, as seen in Eq. 6.77: σ_{3S} can be derived from σ_{1S} and vice-versa.

In the second order conductivity, the one- and two-photon contributions are intimately related and in the third order conductivity the same is true for the one- and three-photon contributions. Going further, in the n-th order conductivity, the one- and n-photon contributions are expected to be related in the same way. In general terms, which contributions must be linked by overall permutation symmetry can be seen by inspection of the possible resonances involving the available optical frequencies. For instance, at fourth order one- and four-photon contributions are interconnected ($\omega_{1234} = \omega_1 + \omega_2 + \omega_3 + \omega_4$), as well as the two- and the three-photon contributions ($-\omega_1 - \omega_2 = -\omega_{1234} + \omega_3 + \omega_4$). This matching of optical resonances can be carried out to arbitrary order in perturbation theory.

6.6 Time-reversal symmetry

As repeated throughout this thesis, evaluation of the nonlinear conductivity of a crystal demands only a knowledge of the FBZ, the band structure and the non-abelian Berry connection for its derivation. These objects define the electronic system, as far as the nonlinear optical properties are concerned. It matters then, to express the consequences of time-reversal symmetry not in terms of the properties of Bloch functions, which can be found in any standard textbook on the application of group theory to solids [88], but instead on the conditions it poses on $\epsilon_{\mathbf{k}a}$ and $\mathcal{A}^{\alpha}_{\mathbf{k}ab}$,

$$\epsilon_{-\mathbf{k}a} = \epsilon_{\mathbf{k}a} \quad (6.80)$$

$$\mathcal{A}^{\alpha}_{-\mathbf{k}ab} = (\mathcal{A}^{\alpha}_{\mathbf{k}ab})^* = \mathcal{A}^{\alpha}_{\mathbf{k}ba} \quad (6.81)$$

These relations form the statement of time-reversal symmetry that will be used here. The condition on the non-abelian Berry connection needs to be satisfied only

for a specific choice of gauge and it implies that the diagonal matrix elements, the abelian Berry connection $\mathcal{A}_{\mathbf{k}aa}$, must be even in $\mathbf{k}$. This, in turn, means

$$\mathcal{F}^{\beta\alpha}_{-\mathbf{k}a} = -\mathcal{F}^{\beta\alpha}_{\mathbf{k}a} \tag{6.82}$$

The Berry curvature is gauge invariant, so this property must always be verified for any time-reversal symmetric system.

These equations are to be contrasted with inversion symmetry, which determines that

$$\epsilon_{-\mathbf{k}a} = \epsilon_{\mathbf{k}a} \tag{6.83}$$

$$\mathcal{A}^{\alpha}_{-\mathbf{k}ab} = -\mathcal{A}^{\alpha}_{\mathbf{k}ab} \tag{6.84}$$

where, again, the condition on the non-abelian Berry connection is valid for some choice of phase factors on the Bloch functions. Independent of these, the Berry curvature is even in $\mathbf{k}$ as a result,

$$\mathcal{F}^{\beta\alpha}_{-\mathbf{k}a} = \mathcal{F}^{\beta\alpha}_{\mathbf{k}a} \tag{6.85}$$

Notice that if both inversion and time-reversal symmetry are present, as is the case with monolayer graphene, then the Berry curvature must be zero. The same cannot be said for the abelian Berry connection, because Eqs. 6.81 and 6.84 may be satisfied for different phases (choices of gauge) in the Bloch function. Inversion and time-reversal symmetry reflect similarly on the band structure (Eqs. 6.80 and 6.83).

As an example, consider the dispersion relation derived from the nearest neighbour tight binding model of monolayer graphene (Section 3.3.2),

$$\epsilon_{\mathbf{k}c} = +t|\Phi(\mathbf{k})| \qquad \epsilon_{\mathbf{k}v} = -t|\Phi(\mathbf{k})| \tag{6.86}$$

with

$$\Phi(\mathbf{k}) = |\Phi(\mathbf{k})|e^{-i\theta} = e^{i\mathbf{k}\cdot\boldsymbol{\delta}_1} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_2} + e^{i\mathbf{k}\cdot\boldsymbol{\delta}_3} \tag{6.87}$$

The energies are even in $\mathbf{k}$, as expected from inversion or time-reversal symmetry.

The non-abelian Berry connection is

$$\mathcal{A}^{\alpha}_{\mathbf{k}cc} = \mathcal{A}^{\alpha}_{\mathbf{k}vv} = -\frac{1}{2}\,\partial^{\alpha}\theta \qquad \mathcal{A}^{\alpha}_{\mathbf{k}cv} = \mathcal{A}^{\alpha}_{\mathbf{k}vc} = \frac{1}{2}\,\partial^{\alpha}\theta \tag{6.88}$$

It can be read from Eq. 6.87 that θ is an odd function of $\mathbf{k}$. Hence, the non-abelian Berry connection in Eqs. 6.88 respects the time-reversal symmetry condition in Eq. 6.81. The inversion symmetry property is, however, not manifest in Eqs. 6.88 and requires a different choice of phases for the eigenbasis in Eqs 3.54.

This analysis involved the tight binding model. Care should be taken when trying to verify these symmetries in continuum Hamiltonians, such as the Dirac Hamiltonian in Eq. 3.60 or in Eq. 3.65, since the change in sign of the Bloch vector refers to the center of the FBZ. In the case of graphene, this means that in order to check Eq. 6.81 or Eq. 6.84 with low energy descriptions, it is required, in evaluating the left hand side of these equations, to not just consider a shift $-\mathbf{q}$ relative to the Dirac point, but also exchanging valleys.

The impact of time-reversal symmetry on the nonlinear optical response can be understood with the use of the resonance-based analysis of Chapter 5. The most direct manifestation of time-reversal symmetry is in the Fermi surface contributions, where several of the F_X integrals vanish if the system is time-reversal symmetric. Due to the parity of the band structure (Eq. 6.80) and the Berry curvature (Eq. 6.82):

$$F_B^{\beta\alpha_1} = 0 \tag{6.89}$$
$$F_A^{\beta\alpha_1\alpha_2} = 0 \tag{6.90}$$
$$F_B^{\beta\alpha_1\alpha_2\alpha_3} = 0 \tag{6.91}$$

These results cover the optical response from linear to third order, but, upon inspection of the integrands, the pattern is clear. Depending on whether an odd or even order nonlinear conductivity is considered, the Fermi surface contribution is solely dictated by the F_A or F_B integrals, respectively[11].

The implications for the $\mathcal{H}_j$ and $\mathcal{I}_j$ integrals are less trivial to demonstrate, but for the most part translate into the following: for time-reversal symmetric systems, these integrals are real. This can be made evident by manipulation of the integrands, using Eqs. 6.80 and 6.81 to combine the values at $\mathbf{k}$ and $-\mathbf{k}$. Take $\mathcal{I}_1$ in linear order as an instructive example,

$$\begin{aligned}
\mathcal{I}_1^{\beta\alpha_1}(\omega_1) &= \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\, \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right) \\
&= \frac{1}{2} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\, \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right) \\
&\quad + \frac{1}{2} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \mathcal{A}^{\beta}_{-\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{-\mathbf{k}ab}\, \Delta f_{-\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{-\mathbf{k}ab}\right) \\
&= \frac{1}{2} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \left(\mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab} + \mathcal{A}^{\beta}_{-\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{-\mathbf{k}ab}\right) \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right) \\
&= \frac{1}{2} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \left(\mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab} + \left(\mathcal{A}^{\beta}_{\mathbf{k}ba}\right)^* \left(\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\right)^*\right) \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right) \\
&= \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \mathrm{Re}\{\mathcal{A}^{\beta}_{\mathbf{k}ba}\, \mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\}\, \Delta f_{\mathbf{k}ba}\, \delta\left(\hbar\omega_1 - \Delta\epsilon_{\mathbf{k}ab}\right)
\end{aligned} \tag{6.92}$$

The exact same procedure works for the $\mathcal{H}_1$ in Eq. 5.50. In general, it is found that the real part of products of Berry connections, or generalized derivatives of Berry connections, appears in the integrands at odd orders in perturbation theory, when subjected to these manipulations, while the imaginary part is relevant at even orders in perturbation theory. For instance, the $\mathcal{H}_j$ integrals at second order (that follow from Eqs. 5.28 and 5.29) can be rewritten as

[11] On the basis that there is not some other F_C integral that appears at higher orders; the derivations have not been carried out by the author beyond third order.

$$\mathcal{H}_1^{\beta\alpha_1\alpha_2}(\bar{\omega}) = \frac{1}{\pi}\fint \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathrm{Im}\{\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\}\,(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab})}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{6.93}$$

$$\mathcal{H}_2^{\beta\alpha_1\alpha_2}(\bar{\omega}) = \frac{1}{\pi}\fint \frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathrm{Im}\{\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab;\alpha_2}\}}{\hbar\bar{\omega}-\Delta\epsilon_{\mathbf{k}ab}}\,\Delta f_{\mathbf{k}ba} \tag{6.94}$$

Either way, the integrals themselves are always real in the presence of time-reversal symmetry. A version of the formulae of Section 5.2 that incorporates time-reversal symmetry can be found in [91], where the previous manipulations were done on every integral.

There is an entire new set of tensor identities that results from these manipulations. Just observe that the $\mathcal{H}_1$ in Eq. 6.93 obeys

$$\mathcal{H}_1^{\beta\alpha_1\alpha_2}(\omega) = -\mathcal{H}_1^{\alpha_1\beta\alpha_2}(\omega) \tag{6.95}$$

and vanishes when considering diagonal tensor elements. All the integral identities resulting from time-reversal symmetry, ranging from linear to third order, are listed in Appendix G.

As briefly discussed in Section 5.3, if the $\mathcal{I}_j$ and $\mathcal{H}_j$ integrals are real, then

$$\mathrm{Re}\{\sigma^{\beta\alpha_1\ldots\alpha_n}\} = \sigma_{\mathcal{I}}^{\beta\alpha_1\ldots\alpha_n} \tag{6.96}$$

$$i\,\mathrm{Im}\{\sigma^{\beta\alpha_1\ldots\alpha_n}\} = \sigma_{\mathcal{H}}^{\beta\alpha_1\ldots\alpha_n} \tag{6.97}$$

The correspondence between $\sigma_{\mathcal{I}}$ and the real part of the nonlinear conductivity, and likewise between $\sigma_{\mathcal{H}}$ and the imaginary part, has significant consequences.

The portion of the conductivity that is properly described by the electronic properties near specific regions of the FBZ, where resonances occur, is $\sigma_{\mathcal{I}}$, not necessarily the real part. The portion of the conductivity that is calculable via a series of "Fermi golden rule"-type computations is $\sigma_{\mathcal{I}}$, not necessarily the real part. The portion of the conductivity that is responsible for optical absorption, as seen in our previous discussion on overall permutation symmetry (Section 6.4), is $\sigma_{\mathcal{I}}$, not necessarily the real part.

If time-reversal symmetry is present, we can expect the real part of the nonlinear optical conductivity to be comprised of a series of Heaviside step functions and to vanish when all optical frequencies of relevance fall below the band gap. The real part is in this case responsible for optical absorption and, for an insulator, there can be no optical losses in that region of the spectrum. We equally expect divergences and a more complicated behavior from the imaginary part. The imaginary part is, generally, nonzero below the band gap and it is from it that the (nonlinear) refractive index is calculated.

But if time-reversal symmetry is broken, the Hilbert transforms $\mathcal{H}_j$ will appear in both the real and imaginary parts of the conductivity. As will the $\mathcal{I}_j$ integrals. Steps appear in the imaginary part and the real part of the nonlinear optical conductivity is no longer zero when all optical frequencies lie below the band gap. The crystal is still transparent for those optical frequencies, of course, and no optical losses are to be expected. It is important to note that the condition of *no optical losses does*

not imply that the real part of the conductivity must be zero below the band gap. For example, the transverse linear conductivity can be purely imaginary, while still respecting the anti-hermiticity condition (Eq. 6.17) and having no energy transferred into the medium. In fact, the existence of a real part of the nonlinear conductivity in the transparency region of the spectrum could be seen as a signature of broken time-reversal symmetry.

Chapter 7

Conclusions

Throughout this thesis, our attention centered around the nonlinear optical conductivities of crystalline solids. The mathematical structure of these response functions was examined in detail, clarified through reworkings of known and discovered formulas, and theoretical edifices were raised to facilitate their computation.

The discussion opened, in Chapter 1, with the hypothesis that optical fields propagating through a medium were sufficiently weak for its constitutive relation to be properly captured by a series expansion. The coefficients in this series expansion are the linear and nonlinear optical conductivities of the medium (Eqs. 1.8 and 1.9). An historical overview of perturbative nonlinear optics then helped set the work of this thesis in context. The majority of the succeeding discussion was directed to answering an important practical question: how to derive the nonlinear conductivities of a crystal, supposing we are sufficiently knowledgeable of its electronic properties?

The required knowledge was shown to be the non-abelian Berry connection and the band structure of the material (Chapter 3). This information is retrieved from the eigenvalues and eigenstates of the Hamiltonian $\hat{H}_0$ describing the electron motion in the lattice potential, prior to optical excitation. The Hamiltonian in Eq. 2.35 seems to be a natural starting point, but is potentially misleading. For any actual calculation, effective models with a finite number of bands will be employed and to avoid the pitfalls that pervade perturbative calculations in nonlinear optics, it is best to consider a more general Hamiltonian:

$$\hat{H}_0 = \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a |\psi_{\mathbf{k}a}\rangle \, \epsilon_{\mathbf{k}a} \, \langle\psi_{\mathbf{k}a}| \tag{7.1}$$

Depending on the number of included bands, this Hamiltonian could be the one in Eq. 2.35 or could have been derived from a tight binding description or any other standard theory on the electronic structure of solids. For the specific case of tight binding, more complete expressions for the position matrix elements were given (Eq. 3.17) than those that are typically used in the literature. Ignoring any orbital overlap recovers the known formula for the Berry connection (Eq. 3.20). The case where the sum in Eq. 7.1 contains only two bands received special attention for the mathematical simplicity and physical insight it provides (see Section 3.2 and Chapter 5).

Once the basic ingredients are gathered, a semiclassical treatment based on the density matrix can be used to find the nonlinear conductivity. A perturbation theory developed on the assumption of a linear coupling,

$$\hat{H} = \hat{H}_0 + \hat{\mathcal{O}}^\alpha \, E^\alpha(t) \tag{7.2}$$

gives formulas for any response function relating an observable $\hat{J}^\alpha$ to the classical field $E^\alpha(t)$ (Eqs. 2.28-2.31).

In nonlinear optics, the light-matter coupling is commonly introduced in one of two ways, shown to be equivalent by an unitary transformation. In the length gauge, the observable $\hat{\mathcal{O}}$ is the dipole operator that couples to the electric field, while in the velocity gauge it is the velocity (more accurately, $e\,p^\alpha/m$) that couples to the potential vector. Despite their equivalence, these two choices of gauge have in the past led researchers to contradicting results. The source of these difficulties was pinpointed already in the nineties [48, 49]: band truncation, as well as the use of low energy Hamiltonians, leads to the violation of sum rules connecting the two formulations. Still, the question remained on why the velocity gauge specifically was plagued with artificial infrared divergences.

Here, gauge invariance was analyzed in further detail and the essential conceptual difficulty, that does not seem to have been properly recognized before, was attributed to the nature of the minimal coupling Hamiltonian, for which the perturbation is functionally dependent on $\hat{H}_0$. Evidence for this is the reformulation of the minimal coupling (velocity gauge) method achieved here (Chapter 4) and the derivation of a new set of sum rules (Eq. 4.50) that connect it to the length gauge method and remain valid for finite band models. With this, the decades-long issue of practical differences between the two choices of gauge is definitively solved.

Curiously, the exploration of these two methods led us to the study of two variations, or generalizations, of the density matrix perturbation theory reviewed in Chapter 2 (based on the Hamiltonian in Eq. 7.2).

In the minimal coupling method (Chapter 4), it is necessary to consider couplings beyond linear order,

$$\hat{H} = \hat{H}_0 + \sum_{n=1}^{+\infty} \frac{e^n}{n!} \, \hat{h}^{\alpha_1 \ldots \alpha_n} \, A^{\alpha_1}(t) \ldots A^{\alpha_n}(t) \tag{7.3}$$

The defining advantage of the minimal coupling Hamiltonian is that it preserves translation symmetry. As a result, no derivatives in reciprocal space are found, except those contained in the definition of the $\hat{h}^{\alpha_1 \ldots \alpha_n}$ operators (Eq. 4.9), making it ideal for the numerical evaluation of the nonlinear conductivity. Numerical integration must be performed over the entire FBZ, in order to satisfy sum rules. For tight binding models, it was shown that the computation of $h^{\alpha_1 \ldots \alpha_n}_{\mathbf{k}ab}$ is usually simple and no knowledge of the Berry connection is required. For any model, $\hat{h}$ involves the covariant differentiation of $\hat{H}_0$ alone and is independent of chemical potential, temperature and the optical frequencies involved, which can all be changed without significant increase in computational cost. This makes the minimal coupling algorithm introduced in this thesis a versatile tool for the numerical computation of nonlinear conductivities.

The minimal coupling method has been demonstrated by a calculation of the third order conductivity of monolayer graphene (Section 4.6) in the context of a nearest neighbour tight binding model [56, 84]. At low frequencies, it reproduces previous results in the literature [76], but these were extended to higher optical frequencies [84], beyond the Dirac point approximation, with no added difficulty.

A diagrammatic implementation of our theory has been proposed and used by others [62, 85, 97], namely to describe third harmonic generation in Weyl semimetals [85].

In the length gauge method (Chapter 5), the coupling in the Hamiltonian is linear, but the quantum operator that couples to the optical field is a covariant derivative,

$$\hat{H} = \hat{H}_0 + i\, e\, \hat{D}^\alpha\, E^\alpha(t) \tag{7.4}$$

The advantage of this method is that it handles approximations much better. Band truncation is not an issue and low energy descriptions can be adopted. Its disadvantage is its complexity, which, due to the repeated covariant differentiation of the density matrix, grows very fast with the order of perturbation theory.

To manage this, a resonance-based decomposition of the nonlinear conductivity was proposed, whereby the lengthy expressions derived from the length gauge method were written as a sum of contributions relating to the Fermi surface, one-photon resonances, two-photon resonances, ... In turn, these contributions were written in terms of a minimal set of integrals over the FBZ. For any reader who has consulted the literature on the nonlinear optical response of solids, the value of these results should be clear. Traditionally extensive and tedious derivations of the third order conductivity were reduced to the mechanical evaluation of a small number of integrals. Furthermore, when considering the relaxation-free limit, these integrals constitute a generalization of the Fermi's golden rule methodology to nonlinear order.

The resonance-based analysis has the additional benefit of granting a better understanding of the physics underlying the dispersion of the nonlinear conductivity; universal characteristics are made transparent and specific results for a given system, such as a divergence in the real part or the absence of a jump discontinuity can be derived with ease.

The length gauge method was introduced by Aversa and Sipe [49] and has already received ample validation, being currently the standard method for obtaining the nonlinear conductivity. The proposed methodology involving the Π_j^α integrals was demonstrated with the derivation of the analytical form of the third order conductivity of graphene, as described by the Dirac Hamiltonian. The inclusion of a mass term in the Hamiltonian (gapped graphene) was also addressed (Appendix H). A comparison with numerical results from the minimal coupling method showed perfect agreement, which confirms the validity of the low energy description for the considered range of optical frequencies.

Finally, the symmetry properties of the nonlinear optical response were analyzed (Chapter 6). Starting with a brief review of the reality condition, the nonlinear Kramer-Krönig relations and crystal spatial symmetries, the bulk of the following discussion concerned overall permutation symmetry. Intrinsic permutation symmetry (Eq. 2.32) stems directly from the definition of a nonlinear response function and is responsible for the mandatory symmetrization that follows any derivation of the nonlinear conductivity. Overall permutation symmetry is an extension of intrinsic permutation symmetry, where all tensor indices can now be permuted with appropriate rearrangements of the related optical frequencies. It is a symmetry property of lossless systems, but, by analytical continuation, its validity was extended for any crystal [27].

By superposing the resonance-based decomposition on the statement of overall permutation symmetry, new identities were uncovered. Overall permutation symmetry permitted a clear identification of the dissipative part of the optical response and established unexpected connections between resonant processes involving different numbers of photons. For instance, one-photon and three-photon processes were shown to be interdependent at third order (Eq. 6.77).

The nonlinear optics of crystalline solids lacking time-reversal symmetry is a topic that currently attracts great interest from the research community. In this thesis, the consequences of time-reversal symmetry on the mathematical structure of the nonlinear conductivity were made transparent: some of the Fermi surface integrals vanish identically (namely, the anomalous Hall conductivity) and the other integrals ($\mathcal{I}_j^\alpha$ and $\mathcal{H}_j^\alpha$) are guaranteed to be real. It was proven that the real part and the dissipative part of the optical response are the same in the presence of time-reversal symmetry. It follows that observation of a real part in the transparency region of the spectrum can be interpreted as a signature of broken time-reversal symmetry.

After these various explorations, what we end up with are two powerful and complementary tools for the analysis and derivation of the nonlinear conductivity: the minimal coupling method presents unprecedented flexibility for numerical full band structure computations, while the length gauge method is capable of retrieving analytical expressions from low energy Hamiltonians, by performing a sequence of "Fermi golden rule calculations" and Hilbert transforms.

These tools can now be applied to investigate other systems beyond monolayer graphene and to further improve and test our knowledge on the optical properties of matter. Future research should turn to other two-dimensional crystals and the more complex (and interesting) Van der Waals heterostructures. But there is no need to constrain ourselves to 2D systems; the reader may implement these methods to explore his material of choice. After all, two-dimensional materials were used as an example due to their simplicity, but the methods themselves were independent of the dimensionality d of the crystal. A subject of particular interest for the author would be study of the nonlinear optics of bilayer graphene (with AB stacking). It was the first material discovered with a tunable band gap [98], by application of an external voltage, and it should be interesting to see the role this feature plays in harmonic generation and other nonlinear phenomena.

Graphene itself merits further study. An accurate account of its band structure requires beyond nearest neighbour hoppings [65], that break the electron-hole symmetry of the system. This may prove too complicated for the derivation of analytical formulas, but is easily accommodated by the minimal coupling algorithm described in this thesis.

The impact of time-reversal symmetry in the nonlinear response, discussed at the end of Chapter 6, is to be tested by studying models where this symmetry is absent (e. g. the Haldane model). Tight binding models where orbital overlap is included also deserve careful study.

In general, the flexibility of the minimal coupling method should be taken advantage of. There is no need to limit calculations to optical frequencies near the Fermi level; the full dispersion curve of the conductivity can be computed. Moreover, the recursive nature of the minimal coupling algorithm means higher orders of perturbation theory, and hence higher harmonics, are promptly accessible. Future research efforts shall use this to help cement the usefulness of the approach.

Further formal development of these methods will also be pursued, however.

The resonance-based analysis was constructed on the basis of a two-band crystal and will take further work to extend it to an arbitrary number of bands. The possibility of double-resonances is anticipated to be a possible obstacle in this endeavour.

The Fermi surface contributions have an elegant and universal form in Eqs. 5.9 and 5.10, with frequency independent integrals, but this was not actually proven, merely verified in practice up to third order. The author of this thesis would venture the guess that this is valid at any order, but the manner by which these expressions emerge is non-obvious, their derivation demanding the perfect cancellation of multiple integrals (such as the ones in Eqs. G.2 and G.6 ; see for instance the derivation in Appendix F). An example of a precise but unanswered question is on the character of the Fermi surface integrals at fourth and higher orders: certainly the types observed so far will be present (F_A and F_B), but will there be other Fermi surface integrals? Further thought should be given to these matters, in particular when trying to understand how the formalism in this thesis relates to alternative approaches in the literature, namely Boltzmann equation methods. It should be possible to attain a better comprehension of the validity of the latter by inspection of the Fermi surface contributions in Chapter 5 and their higher order versions.

Another aspect where the insight brought by the resonance-based decomposition may prove fruitful is in the current search for links between topology and nonlinear optical properties [99]. It is fairly clear, by inspection of the integrands in Π_j^α and F_X^α, that geometrical quantities appear naturally in the nonlinear conductivity. But if quantum geometry appears determinant, it is unclear what role topology would play in characterizing the perturbative nonlinear optical properties of crystals. These discussions normally start from the anomalous, topological, contribution found in linear order (Eq. 5.13), a global term that involved the curvature flow over the FBZ. Yet, no higher order analogues were found in the second and third order conductivities (and are unlikely to exist at any order). This puts into question whether topological arguments are relevant for nonlinear effects, when operating within the single particle picture. An interesting example would be the Weyl semimetals, where claims have been made that the unusually large nonlinear optical constants observed have something to do with topology [100, 101]. The work in this thesis should help verify if this is the case.

Finally, there is a major topic that was intentionally left out of this thesis: the singularities of the nonlinear conductivity. In the absence of relaxation, the nonlinear conductivity diverges for extended regions of frequency space. These divergences are physical and are related to, for instance, current injection [5]. This is an effect of practical interest as it allows coherent control of electric currents by light [55]. The expressions in Chapter 5 are the ideal starting point to investigate this singular behavior. Any physical divergence can be explored by means of a Laurent series in γ, where both the integrals $\Pi_j^\alpha(\bar{\omega})$ and the combinations of complex frequencies that multiply them must be expanded. Alternatively, the expansion can be made for $\gamma = 0$ with respect to a real frequency[1]. An array of interesting optical phenomena is captured by frequency components of the conductivity that diverge in the relaxation-free limit and these will be researched under this perspective: electro-optic effects, the optical Kerr effect, two-color current injection, ... Along this line of research, a correction to the recently proposed jerk current [6, 94, 102] has already been made

[1]For example, $\sigma(\omega + \delta, -\omega)$ can be expanded in δ and is expected to diverge for $\delta \to 0^+$.

with use of these methods [7, 8].

We conclude by returning to the approximations made at the very beginning: the electric dipole and independent electron approximations. There is interest (e. g. in plasmonics) in a more general version of the nonlinear conductivity that takes into account spatial dispersion and nonuniform optical fields (at the unit cell level). Unfortunately, the assumption of a spatially independent electric field underlies the Hamiltonians in Eqs. 7.3 and 7.4 and it is still unclear, at the moment, how to extend the analysis of this thesis in this particular direction.

If the electric dipole approximation is reasonable in most commonly encountered scenarios, the same cannot be said for the neglect of electron-electron interactions. Recent publications suggest that the single particle picture may not properly capture the nonlinear optical response of monolayer graphene [103], even though it has found reasonable agreement with experiment [78]. For gapped systems, the influence of excitons on the optical response cannot be disregarded [104]. It is essential to include a proper description of relaxation mechanisms beyond a phenomenological one and to understand the role electron-electron interactions play in nonlinear optics.

Considering the complexity of the nonlinear optical response of crystalline solids, the framework set up in this thesis to treat independent electrons is likely to remain relevant. It is the first step in describing the quantum electron dynamics when coupled to a classical optical field and forms a basis for future theories to build on and compare to. It is the author's hope that this work will prove useful to researchers in the field attempting a better understanding of the light-matter interactions in solids.

Appendices

Appendix A

On time-dependent unitary transformations

As alluded to in Section 1.2, there are two commonly adopted choices of gauge when performing calculations in nonlinear optics. The velocity gauge, which provides the following Hamiltonian

$$\hat{H} = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\mathbf{A}(t)) \tag{A.1}$$

and the length gauge, which gives

$$\hat{H} = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}}) - e\,\hat{\mathbf{r}} \cdot \mathbf{E}(t) \tag{A.2}$$

For each description, the physical states are represented by a set of kets and the observables by Hermitian operators, here labeled by the subscripts V and L to distinguish between the velocity and length gauge choices, respectively. The two methods, their advantages and inherent difficulties are discussed at length in Chapters 2 and 4. Their equivalence can be seen by performing a time-dependent unitary transformation.

The equivalence will be demonstrated here in the context a single particle picture, since this is the language used in this thesis. An assembly of identical electrons, being a fermionic system, normally requires the use of second quantization notation for a rigorous proof, but, as with everything else discussed in this text, it can be shown that the neglect of electron-electron interactions (independent electron approximation) makes the system entirely equivalent to a collection of single particle quantum systems, with the Fermi-Dirac distribution as the equilibrium distribution.

A quantum system consisting of a single particle has its dynamics, prior to the introduction of an coupling with a classical field (Eqs. A.1 and A.2), governed by an Hamiltonian $\hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}})$.

The Hilbert space of the system is generated by some basis $\{|\psi_L\rangle\}$. Acting on the states are operators $\hat{\mathcal{O}}_L$, whose measurements return, on average, $\langle\psi_L|\,\hat{\mathcal{O}}_L\,|\psi_L\rangle$. These are the predictions quantum mechanics is able to make and they are invariant under an simultaneous unitary transformation of the states and operators,

$$\langle\psi_L|\,\hat{\mathcal{O}}_L\,|\psi_L\rangle = \Big(\langle\psi_L|\,\hat{\mathcal{U}}^\dagger(t)\Big)\,\Big(\hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L\,\hat{\mathcal{U}}^\dagger(t)\Big)\,\Big(\hat{\mathcal{U}}(t)\;|\psi_L\rangle\Big) = \langle\psi_V|\,\hat{\mathcal{O}}_V\,|\psi_V\rangle \tag{A.3}$$

where the unitarity of the transformation $\hat{\mathcal{U}}^\dagger(t)\hat{\mathcal{U}}(t) = \hat{\mathcal{U}}(t)\hat{\mathcal{U}}^\dagger(t) = \hat{\mathbb{1}}$ was used in the first passage, followed by the introduction of the definitions

$$|\psi_V\rangle \equiv \hat{\mathcal{U}}(t)\,|\psi_L\rangle \tag{A.4}$$

$$\hat{\mathcal{O}}_V \equiv \hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L\,\hat{\mathcal{U}}^\dagger(t) \tag{A.5}$$

This demonstrates the equivalence of the descriptions of the quantum system that use the states and operators $\{|\psi_L\rangle\,, \hat{\mathcal{O}}_L\}$ and those that use $\{|\psi_V\rangle\,, \hat{\mathcal{O}}_V\}$.

If instead of pure states, mixed states of the quantum system are of interest, then the density operator provides the most convenient route to computing the ensemble average,

$$< \hat{\mathcal{O}}_L >= \mathrm{Tr}\Big(\hat{\mathcal{O}}_L\,\hat{\rho}_L\Big) \tag{A.6}$$

The trace can be evaluated over any basis, $\{|\psi_L\rangle\}$ or $\{|\psi_V\rangle\}$ or any other. Equivalent descriptions are still obtainable by unitary transformations, since

$$\begin{aligned}
\mathrm{Tr}\Big(\hat{\mathcal{O}}_L\,\hat{\rho}_L\Big) &= \mathrm{Tr}\Big(\Big(\hat{\mathcal{U}}^\dagger(t)\hat{\mathcal{U}}(t)\Big)\;\hat{\mathcal{O}}_L\,\Big(\hat{\mathcal{U}}^\dagger(t)\hat{\mathcal{U}}(t)\Big)\,\hat{\rho}_L\,\Big(\hat{\mathcal{U}}^\dagger(t)\hat{\mathcal{U}}(t)\Big)\Big) \\
&= \mathrm{Tr}\Big(\hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L\,\hat{\mathcal{U}}^\dagger(t)\,\hat{\mathcal{U}}(t)\,\hat{\rho}_L\,\hat{\mathcal{U}}^\dagger(t)\,\Big(\hat{\mathcal{U}}(t)\hat{\mathcal{U}}^\dagger(t)\Big)\Big) \\
&= \mathrm{Tr}\Big(\Big(\hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L\,\hat{\mathcal{U}}^\dagger(t)\Big)\,\Big(\hat{\mathcal{U}}(t)\,\hat{\rho}_L\,\hat{\mathcal{U}}^\dagger(t)\Big)\Big) \\
&= \mathrm{Tr}\Big(\hat{\mathcal{O}}_V\,\hat{\rho}_V\Big)
\end{aligned} \tag{A.7}$$

where both the unitarity of the transformation and the cyclic property of the trace were used.

Consider the case of an unitary transformation with the form,

$$\hat{\mathcal{U}}(t) = e^{-i\,e\,\mathbf{A}(t)\cdot\hat{\mathbf{r}}/\hbar} \tag{A.8}$$

with $\mathbf{A}(t)$ as a classical vector field.

In this case, the momentum operator transforms as,

$$\hat{\mathbf{p}}_L \to \hat{\mathbf{p}}_V = \hat{\mathcal{U}}(t)\,\hat{\mathbf{p}}_L\,\hat{\mathcal{U}}^\dagger(t) = \hat{\mathbf{p}}_L + e\,\mathbf{A}(t) \tag{A.9}$$

and, more generally,

$$\hat{\mathcal{O}}_L(\hat{\mathbf{r}},\hat{\mathbf{p}}) \to \hat{\mathcal{O}}_V(\hat{\mathbf{r}},\hat{\mathbf{p}}) = \hat{\mathcal{U}}(t)\,\hat{\mathcal{O}}_L(\hat{\mathbf{r}},\hat{\mathbf{p}})\,\hat{\mathcal{U}}^\dagger(t) = \hat{\mathcal{O}}_L(\hat{\mathbf{r}},\hat{\mathbf{p}} + e\,\mathbf{A}(t)) \tag{A.10}$$

Knowing how to translate operators and states at any given time between the two descriptions, we are left with the question: what is the Hamiltonian that describes the dynamics of the transformed space of states? A first guess could be

$$\hat{H}_L(\hat{\mathbf{r}},\hat{\mathbf{p}}) \to \hat{H}_V(\hat{\mathbf{r}},\hat{\mathbf{p}}) \stackrel{?}{=} \hat{\mathcal{U}}(t)\,\hat{H}_L(\hat{\mathbf{r}},\hat{\mathbf{p}})\,\hat{\mathcal{U}}^\dagger(t) = \hat{H}_L(\hat{\mathbf{r}},\hat{\mathbf{p}} + e\,\mathbf{A}(t)) \tag{A.11}$$

but this is incorrect. The situation is more complicated because the unitary transformation is time-dependent: the states evolve in time not only due to the action of

the original time evolution operator $\hat{U}(t) = e^{-i\,\hat{H}_L\,t/\hbar}$, but also due to the mapping itself.

The true definition of the Hamiltonian is in the time-dependent Schödinger equation[1], as the generator of the time evolution of physical states

$$\hat{H}_V(\hat{\mathbf{r}}, \hat{\mathbf{p}})\,|\psi_V(t)\rangle \equiv i\,\hbar\,(\partial_t\,|\psi_V(t)\rangle) \tag{A.12}$$

From here, the general form of an Hamiltonian subjected to a time-dependent unitary transformation can be derived,

$$\begin{aligned} i\,\hbar\,(\partial_t\,|\psi_V(t)\rangle) &= i\,\hbar\,\Big(\partial_t\,\hat{\mathcal{U}}(t)\,|\psi_L(t)\rangle\Big) \\ &= i\,\hbar\,(\partial_t\hat{\mathcal{U}}(t))\,|\psi_L(t)\rangle + i\,\hbar\,\hat{\mathcal{U}}(t)\,(\partial_t\,|\psi_L(t)\rangle) \\ &= i\,\hbar\,(\partial_t\hat{\mathcal{U}}(t))\,|\psi_L(t)\rangle + \hat{\mathcal{U}}(t)\,\hat{H}_L\,|\psi_L(t)\rangle \\ &= i\,\hbar\,(\partial_t\hat{\mathcal{U}}(t))\,\hat{\mathcal{U}}^\dagger(t)\,|\psi_V(t)\rangle + \hat{\mathcal{U}}(t)\,\hat{H}_L\,\hat{\mathcal{U}}^\dagger(t)\,|\psi_V(t)\rangle \end{aligned} \tag{A.13}$$

We conclude,

$$\hat{H}_V = \hat{\mathcal{U}}(t)\,\hat{H}_L\,\hat{\mathcal{U}}^\dagger(t) + i\,\hbar\,(\partial_t\hat{\mathcal{U}}(t))\,\hat{\mathcal{U}}^\dagger(t) \tag{A.14}$$

If we take Eq. A.2,

$$\hat{H}_L = \hat{H}_0 + e\,\hat{\mathbf{r}}\cdot\mathbf{E}(t) \tag{A.15}$$

and apply the unitary transformation in Eq. A.8, we arrive at

$$\begin{aligned} \hat{H}_V(\hat{\mathbf{r}}, \hat{\mathbf{p}}) &= \hat{\mathcal{U}}(t)\,\hat{H}_L(\hat{\mathbf{r}}, \hat{\mathbf{p}})\,\hat{\mathcal{U}}^\dagger(t) + i\,\hbar\,(\partial_t\hat{\mathcal{U}}(t))\,\hat{\mathcal{U}}^\dagger(t) \\ &= \hat{H}_L(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\mathbf{A}(t)) + e\,\hat{\mathbf{r}}\cdot(\partial_t\mathbf{A}(t)) \\ &= \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\mathbf{A}(t)) + e\,\hat{\mathbf{r}}\cdot\mathbf{E}(t) + e\,\hat{\mathbf{r}}\cdot(\partial_t\mathbf{A}(t)) \end{aligned} \tag{A.16}$$

This is the end result for an arbitrary $\mathbf{A}(t)$. If we judiciously pick $\mathbf{A}(t)$ to be such that $-\partial_t\mathbf{A}(t) = \mathbf{E}(t)$ (in other words, if we interpret $\mathbf{A}(t)$ as the vector potential), then

$$\hat{H}_V(\hat{\mathbf{r}}, \hat{\mathbf{p}}) = \hat{H}_0(\hat{\mathbf{r}}, \hat{\mathbf{p}} + e\,\mathbf{A}(t)) \tag{A.17}$$

showing that the two standard ways of introducing an external electric field in the Hamiltonian are indeed equivalent and related by a simple time-dependent unitary transformation.

The results derived here did not depend on the nature of the quantum system that was coupled with the classical field: the reasoning is valid for any $\hat{H}_0$. In this thesis, the quantum system of interest is an electron moving within the periodic potential of a crystal, as represented by either the standard Schrödinger Hamiltonian in Eq. 2.35, or by a finite band model, as in Eq. 4.1. The previous proof on the equivalence of the length and velocity gauge descriptions holds in both cases, although in the latter the position operator is to be understood as the covariant derivative in Eq. 2.50.

[1] As in the rest of this text, the Schrödinger picture of time evolution is adopted.

Appendix B

On the relation between susceptibility and conductivity

This appendix addresses the relationship between two sets of response functions, electric susceptibilities and conductivities, that are of common use in nonlinear optics studies. Each of them provides a complete description of the optical properties of the system, so knowledge of one permits the derivation of the other.

Early developments in nonlinear optics regarded atomic/molecular gases and dielectrics, for which the notion of a macroscopic polarization built out of an average of individual electric dipole moments is intuitive. This led to the susceptibility being the response function traditionally adopted in nonlinear optics, at least in earlier literature.

On the other hand, for many solids, such as doped semiconductors and metals, where DC currents may flow freely, this picture of a density of electric dipole moments is not appropriate. Typically, an effective polarization (and hence a susceptibility) is defined, by considering the quantity whose time derivative gives the electric current [9]. But then, one could argue, a direct treatment of the current feels more natural. Additionally, the electric current is more a fundamental quantity in field theory, due to its connection with symmetry via Noether's theorem. Whenever a constant DC current is injected, the susceptibility diverges, while the conductivity stays finite. For all these reasons, it is the author's conviction that the conductivity provides a more physical and transparent approach to understanding optical properties. In this thesis, where the focus is on the nonlinear optical properties of crystals, conductivities have been almost exclusively used.

Occasionally, the susceptibility has been invoked. For instance, in the discussion on overall permutation symmetry (Section 6.2), where it added useful insight. Also, the conversion between conductivity and susceptibility is often necessary when comparing results with the literature. It is therefore worthwhile to clarify the connection between these two objects.

The linear conductivity is defined by

$$J^{\beta}(t) = \int_{-\infty}^{+\infty} \sigma^{\beta\alpha}(t - t')\, E^{\alpha}(t')\, dt' \tag{B.1}$$

where J is the electric current.

The linear susceptibility is defined by

$$P^{\beta}(t) = \epsilon_0 \int_{-\infty}^{+\infty} \chi^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' \tag{B.2}$$

where P is the electric polarization and ϵ_0 is the vacuum permittivity.

It is known that

$$J^{\beta}(t) = \frac{dP^{\beta}}{dt} \tag{B.3}$$

Replacing Eqs. B.1-B.2 in Eq. B.3,

$$\begin{aligned}\int_{-\infty}^{+\infty} \sigma^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' &= \epsilon_0 \frac{d}{dt} \int_{-\infty}^{+\infty} \chi^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' \\ &= \epsilon_0 \int_{-\infty}^{+\infty} \left(\frac{d}{dt} \chi^{\beta\alpha}(t-t') \right) E^{\alpha}(t')\, dt'\end{aligned} \tag{B.4}$$

This is true for any electric field $E(t)$, which implies that

$$\sigma^{\beta\alpha}(t) = \epsilon_0 \frac{d\chi^{\beta\alpha}}{dt} \tag{B.5}$$

Thus, the conductivity is simply the time derivative of the electric susceptibility. After a Fourier transform is applied, this translates into the following relation between frequency components,

$$\sigma^{\beta\alpha}(\bar{\omega}) = -i\,\bar{\omega}\,\epsilon_0\,\chi^{\beta\alpha}(\bar{\omega}) \tag{B.6}$$

Similarly for nonlinear orders, where the nonlinear response functions are defined by

$$\begin{aligned}J^{\beta\,(n)}(t) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \sigma^{\beta\alpha_1\ldots\alpha_n}(t-t_1,\ldots,t-t_n)\, E^{\alpha_1}(t_1)\ldots E^{\alpha_n}(t_n)\, dt_1\ldots dt_n \\ &= \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi} \ldots \frac{d\omega_n}{2\pi}\; \sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)\, E^{\alpha_1}(\omega_1)\ldots E^{\alpha_n}(\omega_n)\, e^{-i(\bar{\omega}_1+\ldots+\bar{\omega}_n)t}\end{aligned} \tag{B.7}$$

and

$$\begin{aligned}P^{\beta\,(n)}(t) &= \epsilon_0 \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \chi^{\beta\alpha_1\ldots\alpha_n}(t-t_1,\ldots,t-t_n)\, E^{\alpha_1}(t_1)\ldots E^{\alpha_n}(t_n)\, dt_1\ldots dt_n \\ &= \epsilon_0 \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} \frac{d\omega_1}{2\pi} \ldots \frac{d\omega_n}{2\pi}\; \chi^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)\, E^{\alpha_1}(\omega_1)\ldots E^{\alpha_n}(\omega_n)\, e^{-i(\bar{\omega}_1+\ldots+\bar{\omega}_n)t}\end{aligned} \tag{B.8}$$

The relations between nonlinear susceptibilities and conductivities can be read from Eqs. B.3, B.7 and B.8,

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(t) = \epsilon_0 \frac{d\chi^{\beta\alpha_1\ldots\alpha_n}}{dt} \tag{B.9}$$

and

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) = -i\,(\bar{\omega}_1+\ldots+\bar{\omega}_n)\,\epsilon_0\,\chi^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n) \tag{B.10}$$

It should be possible to verify these relations with the expressions provided by perturbation theory in Section 2.2. To keep the setting as general as possible, and hence avoid extraneous details, we consider that the Hamiltonian governing the system is $\hat{H}(t) = \hat{H}_0 + \hat{\mathcal{V}}(t)$, where $\hat{H}_0$ is well-understood and the second term is a perturbation of the form $\hat{\mathcal{V}}(t) = \mathcal{O}^\alpha\,E^\alpha(t)$. These were the basic assumptions from which the perturbative treatment of Sections 2.1-2.2 was developed. The formulas for linear and nonlinear response functions were given in Eqs. 2.28-2.29. Alternatively, they can evaluated with Eqs. 2.30 and 2.31.

Taking Eq. 2.30 for $n = 1$, we obtain the standard Kubo's formula for the linear conductivity,

$$\sigma^{\beta\alpha}(t) = \left(\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{\mathcal{O}}^\alpha_I(-t),\hat{J}^\beta\right]\right)\,\Theta(t) \tag{B.11}$$

where $\hat{\rho}_0$ is the equilibrium distribution and $\hat{\mathcal{O}}_I(t)$ is the result of evolving the operator $\hat{\mathcal{O}}$ in the interaction picture.

Similarly, for the susceptibility,

$$\epsilon_0\,\chi^{\beta\alpha}(t) = \left(\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{\mathcal{O}}^\alpha_I(-t),\hat{P}^\beta\right]\right)\,\Theta(t) \tag{B.12}$$

The time-dependence can, somewhat artificially, be transferred to the polarization by evolving every operator for an interval of time t in the interaction picture (this corresponds to performing the unitary transformation $\hat{U}_0(t) = e^{-i\hat{H}_0 t/\hbar}$),

$$\epsilon_0\,\chi^{\beta\alpha}(t) = \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^\beta_I(t),\hat{\mathcal{O}}^\alpha\right]\right)\,\Theta(t) \tag{B.13}$$

In attempting to demonstrate Eq. B.5, the following manipulations could be attempted on Eq. B.11

$$\begin{aligned}
\sigma^{\beta\alpha}(t) &= \left(\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{\mathcal{O}}^\alpha_I(-t),\hat{J}^\beta\right]\right)\,\Theta(t)\\
&= \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{J}^\beta_I(t),\hat{\mathcal{O}}^\alpha\right]\right)\,\Theta(t)\\
&= \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\frac{d\hat{P}^\beta_I}{dt},\hat{\mathcal{O}}^\alpha\right]\right)\,\Theta(t)\\
&= \frac{d}{dt}\left(\left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^\beta_I(t),\hat{\mathcal{O}}^\alpha\right]\right)\right)\,\Theta(t)
\end{aligned} \tag{B.14}$$

but there are several difficulties with this result. It is, in fact, in contradiction with Eq. B.5.

To recognize this, note that the susceptibility (Eq. B.12) has the form,

$$\chi^{\beta\alpha}(t) = X^{\beta\alpha}(t)\,\Theta(t) \tag{B.15}$$

where $X^{\beta\alpha}(t)$ is a well-behaved function that can be made differentiable at $t = 0$. A direct differentiation of this expression should give the conductivity,

$$\sigma^{\beta\alpha}(t) = \epsilon_0 \frac{d\chi^{\beta\alpha}}{dt} = \epsilon_0 \frac{dX^{\beta\alpha}}{dt}\,\Theta(t) + \epsilon_0\, X^{\beta\alpha}(0)\,\delta(t) \tag{B.16}$$

but the second term in Eq. B.16 is absent from Eq. B.14.

Indeed, it is unclear how could such a Dirac delta function even appear from an application of Kubo's formula. One might feel that maybe the previous derivation, which involves the derivative of a step function and distributions, is to be questioned, but a more careful and rigorous analysis shows that Eq. B.16 holds:

$$\begin{aligned} J^{\beta}(t) &= \epsilon_0 \frac{d}{dt} \int_{-\infty}^{+\infty} \chi^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' \\ &= \epsilon_0 \frac{d}{dt} \int_{-\infty}^{t} X^{\beta\alpha}(t-t')\, E^{\alpha}(t')\, dt' \\ &= \epsilon_0 \int_{-\infty}^{t} \left(\frac{d}{dt} X^{\beta\alpha}(t-t') \right) E^{\alpha}(t')\, dt' + \epsilon_0\, X^{\beta\alpha}(0)\, E^{\alpha}(t) \\ &= \epsilon_0 \int_{-\infty}^{+\infty} \left(\frac{dX^{\beta\alpha}}{dt}\,\Theta(t-t') + X^{\beta\alpha}(0)\,\delta(t-t') \right) E^{\alpha}(t')\, dt' \end{aligned} \tag{B.17}$$

As a consequence, we recover

$$\sigma^{\beta\alpha}(t) = \epsilon_0 \frac{d\chi^{\beta\alpha}}{dt} = \epsilon_0 \frac{dX^{\beta\alpha}}{dt}\,\Theta(t) + \epsilon_0\, X^{\beta\alpha}(0)\,\delta(t) \tag{B.18}$$

Having the relation between the susceptibility and the conductivity on firm footing, we are left to question our derivation of the linear conductivity (Eq. B.14).

In fact, our misstep relates to a fundamental question: what operator do we define as the electric current, whose time evolution we are interested in following? Perhaps, the operator as it was defined previous to applying the perturbation?

$$\hat{J}^{\beta} = \left(-\frac{i}{\hbar} \right) \left[\hat{P}^{\beta}, \hat{H}_0 \right] \tag{B.19}$$

This definition was the one used to derive Eq. B.14. But, as soon as the perturbation is turned on, this current is no longer related to the electric polarization by a time derivative. If we maintain that the electric current *is defined as* $-e\,\hat{v}^{\beta} = \partial \hat{P}^{\beta}/\partial t$ (or equivalently that the velocity is always defined as the time derivative of the position), then, upon applying the perturbation, the current operator becomes explicitly time-dependent and is given by,

$$\hat{J}^{\beta}(t) = \left(-\frac{i}{\hbar} \right) \left[\hat{P}^{\beta}, \hat{H}(t) \right] = \left(-\frac{i}{\hbar} \right) \left[\hat{P}^{\beta}, \hat{H}_0 \right] + \left(-\frac{i}{\hbar} \right) \left[\hat{P}^{\beta}, \hat{\mathcal{O}}^{\alpha} \right] E^{\alpha}(t) \tag{B.20}$$

This is the correct expression for the current, the one that preserves the relation between $\hat{J}$ and $\hat{P}$. It reduces to Eq. B.19 only when $\hat{P}$ and $\hat{\mathcal{O}}$ commute.

This statement is somewhat obvious, but it implies that when performing perturbation theory one cannot blindly use Kubo's formula for the current. A proper development of perturbation theory involves an expansion in the current matrix

elements as well as in the density matrix, in a way directly analogous to what is observed in a minimal coupling approach [56].

The ensemble average of Eq. B.20 gives

$$J^{\beta}(t) = \mathrm{Tr}\left(\hat{J}^{\beta}(t)\,\hat{\rho}(t)\right) = \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\left[\hat{P}^{\beta},\hat{H}_0\right]\hat{\rho}(t)\right) + \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\left[\hat{P}^{\beta},\hat{\mathcal{O}}^{\alpha}\right]\hat{\rho}(t)\right)\,E^{\alpha}(t) \tag{B.21}$$

In linear order,

$$\begin{aligned} J^{\beta(1)}(t) = &\left(-\frac{i}{\hbar}\right)^2\int_{-\infty}^{t}\mathrm{Tr}\left(\left[\hat{P}^{\beta},\hat{H}_0\right]\left[\hat{\mathcal{O}}^{\alpha}_I(t'-t),\hat{\rho}_0\right]\right)E^{\alpha}(t')\,dt' \\ &+\left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\left[\hat{P}^{\beta},\hat{\mathcal{O}}^{\alpha}\right]\hat{\rho}_0\right)E^{\alpha}(t) \end{aligned} \tag{B.22}$$

corresponding to a conductivity

$$\begin{aligned} \sigma^{\beta\alpha}(t) &= \left(-\frac{i}{\hbar}\right)^2\mathrm{Tr}\left(\hat{\rho}_0\,\left[\left[\hat{P}^{\beta},\hat{H}_0\right],\hat{\mathcal{O}}^{\alpha}_I(-t)\right]\right)\Theta(t) + \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^{\beta},\hat{\mathcal{O}}^{\alpha}\right]\right)\delta(t) \\ &= \left(-\frac{i}{\hbar}\right)^2\mathrm{Tr}\left(\hat{\rho}_0\,\left[\left[\hat{P}^{\beta}_I(t),\hat{H}_0\right],\hat{\mathcal{O}}^{\alpha}\right]\right)\Theta(t) + \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^{\beta}_I(0),\hat{\mathcal{O}}^{\alpha}\right]\right)\delta(t) \\ &= \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\frac{d\hat{P}^{\beta}_I(t)}{dt},\hat{\mathcal{O}}^{\alpha}\right]\right)\Theta(t) + \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^{\beta}_I(0),\hat{\mathcal{O}}^{\alpha}\right]\right)\delta(t) \\ &= \frac{d}{dt}\left(\left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^{\beta}_I(t),\hat{\mathcal{O}}^{\alpha}\right]\right)\right)\Theta(t) + \left(-\frac{i}{\hbar}\right)\mathrm{Tr}\left(\hat{\rho}_0\,\left[\hat{P}^{\beta}_I(0),\hat{\mathcal{O}}^{\alpha}\right]\right)\delta(t) \\ &= \epsilon_0\,\frac{dX^{\beta\alpha}}{dt}\,\Theta(t) + \epsilon_0\,X^{\beta\alpha}(0)\,\delta(t) = \epsilon_0\,\frac{d\chi^{\beta\alpha}}{dt} \end{aligned} \tag{B.23}$$

that is indeed the time derivative of the susceptibility.

It is now clear that, once the proper definitions of the current (Eq. B.20) and respective conductivity (Eq. B.23) are adopted, the relationship between conductivity and susceptibility is as straightforward as expected.

The equivalent relation in the frequency domain, Eq. B.6, can also be demonstrated explicitly[1]:

$$\begin{aligned} -i\,\bar{\omega}\,\chi^{\beta\alpha}(\bar{\omega}) &= -i\,\bar{\omega}\sum_{a,b}P^{\beta}_{ba}\,\frac{[\mathcal{O}^{\alpha},\rho_0]_{ab}}{\hbar\bar{\omega}-\Delta\epsilon_{ab}} \\ &= -\frac{i}{\hbar}\sum_{a,b}\frac{\hbar\bar{\omega}\,P^{\beta}_{ba}}{\hbar\bar{\omega}+\Delta\epsilon_{ba}}\,[\mathcal{O}^{\alpha},\rho_0]_{ab} \\ &= -\frac{i}{\hbar}\sum_{a,b}\left(1-\frac{\Delta\epsilon_{ba}}{\hbar\bar{\omega}+\Delta\epsilon_{ba}}\right)\,P^{\beta}_{ba}\,[\mathcal{O}^{\alpha},\rho_0]_{ab} \end{aligned} \tag{B.24}$$

[1] If $\hat{H}_0$ is the Hamiltonian for an electron moving in a crystalline potential (Eq. 2.35), then there is a continuum of energies and the sum over states includes an integration over the FBZ.

$$
\begin{aligned}
&= -\frac{i}{\hbar}\sum_{a,b}\frac{P^{\beta}_{ba}\,\Delta\epsilon_{ab}}{\hbar\bar{\omega}+\Delta\epsilon_{ba}}\,[\mathcal{O}^{\alpha},\rho_0]_{ab} - \frac{i}{\hbar}\sum_{a,b}P^{\beta}_{ba}\,[\mathcal{O}^{\alpha},\rho_0]_{ab} \\
&= \sum_{a,b}J^{\beta(0)}_{ba}\,\frac{[\mathcal{O}^{\alpha},\rho_0]_{ab}}{\hbar\bar{\omega}-\Delta\epsilon_{ab}} + \left(-\frac{i}{\hbar}\right)\,\mathrm{Tr}\left(\hat{P}^{\beta}\left[\hat{\mathcal{O}}^{\alpha},\hat{\rho}_0\right]\right) \\
&= \mathrm{Tr}\left(\hat{J}^{\beta(0)}\,\hat{\rho}^{\alpha}(\bar{\omega})\right) + \left(-\frac{i}{\hbar}\right)\,\mathrm{Tr}\left(\left[\hat{P}^{\beta},\hat{\mathcal{O}}^{\alpha}\right]\hat{\rho}_0\right) \\
&= \sigma^{\beta\alpha}(\bar{\omega})
\end{aligned}
\tag{B.25}
$$

The previous arguments can be recreated for nonlinear orders to demonstrate Eqs. B.9 and B.10: the formulas are lengthier, but the fundamental concepts and algebraic manipulations are essentially the same.

The identification of an extra term in the definition of the electric current is a direct consequence of the definition of the velocity of a particle whose dynamics are governed by the Hamiltonian $\hat{H}(t) = \hat{H}_0 + \hat{\mathcal{O}}^{\alpha}\,E^{\alpha}(t)$. This extra term vanishes if $\hat{P}$ and $\hat{\mathcal{O}}$ commute, which is expected to be case when $\hat{\mathcal{O}}^{\alpha} = e\,\hat{r}^{\alpha} = -\hat{P}^{\alpha}$. In Appendix D, it is argued that this may not be true, however, for finite band models.

In finite band models of crystals, the position operator is a covariant derivative, which poses some problems in attempting to compute the electric susceptibility, linear or nonlinear, via standard perturbation theory. As discussed in Section 2.3, the position operator does not have well defined matrix elements in the Bloch basis, but can be accommodated by being placed inside commutators whose matrix elements are well defined. This works fine when performing the perturbation theory for the electric current and deriving conductivities, but fails on treating the electric polarization, since in its ensemble average,

$$
P^{\beta} = \mathrm{Tr}\left(\hat{P}^{\beta}\,\hat{\rho}(t)\right) = -e\;\mathrm{Tr}\left(\hat{r}^{\beta}\,\hat{\rho}(t)\right) \stackrel{?}{=} -i\,e\;\mathrm{Tr}\left(\hat{D}^{\beta}\,\hat{\rho}(t)\right) \tag{B.26}
$$

the position operator does not occur inside a commutator. This is the primary reason why a direct calculation of the electric susceptibility in crystals is not possible in the context of a single particle picture.

To conclude, it is emphasized that our observables of interest, $\hat{P}$ and $\hat{J}$, could be replaced by any two operators related by a time derivative and the structure of the reasoning presented in this appendix would, for the most part, remain unaltered. The apparent contradiction between Eqs. B.12, B.14 and B.5 serves as a cautionary tale when developing a perturbation theory for observables that are inherently defined as time derivatives of other physical quantities. In particular, Kubo's formula cannot, in these situations, be directly applied to compute the respective linear and nonlinear response functions.

Appendix C

Commutative covariant derivatives

Throughout a significant portion of this work, it was implicitly assumed that position operators commute. This may seem a trivial statement, derived from the basics of quantum mechanics, but commutation relations tend to be broken upon band truncation and when a finite band model is used to describe a system, defined perhaps by the specification of a set of Bloch states $\psi_{\mathbf{k}a}$ and bands $\epsilon_{\mathbf{k}a}$ (Eq. 4.1), there is no reason to assume that the following will be true,

$$\left[\hat{r}^\alpha, \hat{r}^\beta\right] = -\left[\hat{D}^\alpha, \hat{D}^\beta\right] = 0 \tag{C.1}$$

Still, this particular commutation relation can be shown to hold on very general grounds,

$$\begin{aligned}
\left[\hat{D}^\alpha, \hat{D}^\beta\right]_{\mathbf{k}ab} &= -i\,\partial^\alpha \mathcal{A}^\beta_{\mathbf{k}ab} + i\,\partial^\beta \mathcal{A}^\alpha_{\mathbf{k}ab} - \left[\mathcal{A}^\alpha, \mathcal{A}^\beta\right]_{\mathbf{k}ab} \\
&= \partial^\alpha \left(\left\langle u_{\mathbf{k}a}\middle|\partial^\beta u_{\mathbf{k}b}\right\rangle\right) - \partial^\beta \left(\langle u_{\mathbf{k}a}|\partial^\alpha u_{\mathbf{k}b}\rangle\right) \\
&\quad + \sum_c \left(\langle u_{\mathbf{k}a}|\partial^\alpha u_{\mathbf{k}c}\rangle \left\langle u_{\mathbf{k}c}\middle|\partial^\beta u_{\mathbf{k}b}\right\rangle - \left\langle u_{\mathbf{k}a}\middle|\partial^\beta u_{\mathbf{k}c}\right\rangle \langle u_{\mathbf{k}c}|\partial^\alpha u_{\mathbf{k}b}\rangle\right) \\
&= \left\langle \partial^\alpha u_{\mathbf{k}a}\middle|\partial^\beta u_{\mathbf{k}b}\right\rangle - \left\langle \partial^\beta u_{\mathbf{k}a}\middle|\partial^\alpha u_{\mathbf{k}b}\right\rangle \\
&\quad - \sum_c \left(\langle \partial^\alpha u_{\mathbf{k}a}|u_{\mathbf{k}c}\rangle \left\langle u_{\mathbf{k}c}\middle|\partial^\beta u_{\mathbf{k}b}\right\rangle - \left\langle \partial^\beta u_{\mathbf{k}a}\middle|u_{\mathbf{k}c}\right\rangle \langle u_{\mathbf{k}c}|\partial^\alpha u_{\mathbf{k}b}\rangle\right) \\
&= 0
\end{aligned} \tag{C.2}$$

with the last step making use of the closure relation,

$$\sum_c |u_{\mathbf{k}c}\rangle \langle u_{\mathbf{k}c}| = \hat{\mathbb{1}} \tag{C.3}$$

Eq. C.3 is the essential assumption in the construction of the finite-band model that ensures the validity of Eq. C.1. With the commutation of position operators on firm grounds, some of its consequences can be discussed.

A curious and sometimes useful way to rewrite the statement of Eq. C.1 is

$$\left[\hat{D}^\alpha, \hat{\mathcal{A}}^\beta\right]_{\mathbf{k}ab} = \partial^\beta \mathcal{A}^\alpha_{\mathbf{k}ab} \tag{C.4}$$

From Eq. C.1, various other identities can be derived. Expanding the commutator,

$$\partial^{\alpha}\mathcal{A}^{\beta}_{\mathbf{k}ab} - \partial^{\beta}\mathcal{A}^{\alpha}_{\mathbf{k}ab} = i\,\left[\mathcal{A}^{\alpha},\mathcal{A}^{\beta}\right]_{\mathbf{k}ab} \tag{C.5}$$

In the particular case that $a = b$, this identity relates the abelian Berry curvature to the off-diagonal matrix elements of the Berry connection,

$$\mathcal{F}^{\alpha\beta}_{a} \equiv \partial^{\alpha}\mathcal{A}^{\beta}_{\mathbf{k}aa} - \partial^{\beta}\mathcal{A}^{\alpha}_{\mathbf{k}aa} = i\left[\mathcal{A}^{\alpha},\mathcal{A}^{\beta}\right]_{\mathbf{k}aa} = i\sum_{c\neq a}\left(\mathcal{A}^{\alpha}_{\mathbf{k}ac}\mathcal{A}^{\beta}_{\mathbf{k}ca} - \mathcal{A}^{\beta}_{\mathbf{k}ac}\mathcal{A}^{\alpha}_{\mathbf{k}ca}\right) \tag{C.6}$$

For a two-band model, Eq. C.5 with $a \neq b$ relates "generalized derivatives" of Berry connections,

$$\mathcal{A}^{\beta}_{\mathbf{k}ab;\alpha} = \mathcal{A}^{\alpha}_{\mathbf{k}ab;\beta} \tag{C.7}$$

A more sophisticated identity on second order "generalized derivatives" can also be derived,

$$\mathcal{A}^{\beta}_{\mathbf{k}ab;\alpha_1\alpha_2} - \mathcal{A}^{\beta}_{\mathbf{k}ab;\alpha_2\alpha_1} = 2\left(\mathcal{A}^{\alpha_1}_{\mathbf{k}ba}\mathcal{A}^{\beta}_{\mathbf{k}ab}\mathcal{A}^{\alpha_2}_{\mathbf{k}ab} - \mathcal{A}^{\alpha_2}_{\mathbf{k}ba}\mathcal{A}^{\beta}_{\mathbf{k}ab}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\right) \tag{C.8}$$

Eq. C.8 translates into a relation between the Π_4 and Π_5 integrals at third order, which are no longer independent,

$$\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) - \Pi_4^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) = 2\left(\Pi_5^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}) - \Pi_5^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega})\right) \tag{C.9}$$

Appendix D

Noncommutative covariant derivatives

In Appendix C, a set of important identities were derived on the basis that position operators commute. This assumption was behind many derivations in the main text of this thesis and, in particular, the construction of the expressions in Section 5.2. Here, the opposing scenario of noncommutative position operators (covariant derivatives) is studied.

$$\left[\hat{r}^{\alpha}, \hat{r}^{\beta}\right] \neq 0 \tag{D.1}$$

The intent to abandon what seems to be an obvious fact may seem surprising, but notice that band truncation invalidates most commutator relations, by removal of intermediate states, like those in the implicit sum of Eq. D.1. Once more, the difficulty lies in considering a limited number of bands for our system. There is no doubt that position operators commute in their common definition in the Hilbert space of the Hamiltonian in Eq. 2.35 (this is sometimes regarded as a postulate of quantum mechanics), but for a finite band model the position operator is defined as a covariant derivative and

$$[r^{\alpha}, r^{\beta}]_{\mathbf{k}ab} = -[D^{\alpha}, D^{\beta}]_{\mathbf{k}ab} = i\left(\partial^{\alpha}\mathcal{A}^{\beta}_{\mathbf{k}ab} - \partial^{\beta}\mathcal{A}^{\alpha}_{\mathbf{k}ab}\right) + [\mathcal{A}^{\alpha}, \mathcal{A}^{\beta}]_{\mathbf{k}ab} \tag{D.2}$$

there is no guarantee this commutator vanishes.

However, for every model that was analyzed in this thesis, the commutator indeed vanished. It is zero for any tight binding model in the limit of no overlap considered in Section 3.1. In fact, it vanishes whenever a finite matrix of functions of $\mathbf{k}$ is postulated as the parametric Hamiltonian, from which the non-abelian Berry connection is to be derived. The common trait in these models is that the closure relation $\sum_a |u_{\mathbf{k}a}\rangle \langle u_{\mathbf{k}a}| = 1$ is satisfied, from which it can be proved that the commutator must be zero (Appendix C).

The starting point is not always a parametric Hamiltonian though. For instance, band structure and Bloch functions may be obtained numerically by density functional theory. Also, it is always possible that from an original Hilbert space of many bands, two or three are selected as interesting and by band truncation a finite band Hamiltonian (Eq. 4.1) is defined. Orbital overlap may be included in a tight binding model. In any of these cases, there is no means to ensure the commutator in Eq. D.2

vanishes, that is, to ensure covariant derivatives commute in this truncated band space.

Even if these considerations are discarded as irrelevant, there is a second reason to consider the possibility $[D^\alpha, D^\beta] \neq 0$. The statement that the commutator in Eq. D.2 is zero induces a series of identities, listed in Appendix C, that obscure more important symmetries present in the nonlinear conductivity. The most significant example is overall permutation symmetry, which becomes hard to recognize if the resonance-based analysis is performed on top of the perturbative treatment of Section 2.4.1, instead of the version that will be presented here. There are many identities, even linking different integrals (Eq. C.9), that the equations turn out to obey overall permutation symmetry, as they must (see Chapter 6), but in a rather less obvious way (meaning it can be a non-trivial task to prove the symmetry is present).

For these reasons, we recognize the possibility of noncommutative covariant derivatives and extend the previous, standard, length gauge perturbation theory in Section 2.4.1, accordingly. Skimming across this section, it seems there ought to make no difference. But there is one point where commutative covariant derivatives were indeed assumed: in the definition of the velocity,

$$\hat{v}^\beta = \dot{\hat{r}}^\beta = -\frac{i}{\hbar}[\hat{r}^\beta, \hat{H}] = -\frac{i}{\hbar}[\hat{r}^\beta, \hat{H}_0] - \frac{i\,e}{\hbar}[\hat{r}^\beta, \hat{r}^\alpha]\,E^\alpha(t) \tag{D.3}$$

On account of the noncommutativity in Eq. D.1, there is an extra term in the velocity. The situation is similar to that found in the velocity gauge treatment of Section 2.4.2, namely Eq. 2.70, whose analogue is

$$J^\beta(t) = \frac{i\,e}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}(t)\Big) + \frac{i\,e^2}{\hbar}\,\mathrm{Tr}\big([\hat{r}^\beta, \hat{r}^\alpha]\,\hat{\rho}(t)\big)E^\alpha(t) \tag{D.4}$$

This is the form of the electric current when position operators are not assumed to commute. Expanded in the optical fields, it gives, in linear order,

$$J^{\beta(1)}(t) = \frac{i\,e}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}^{(1)}(t)\Big) + \frac{i\,e^2}{\hbar}\,\mathrm{Tr}\big([\hat{r}^\beta, \hat{r}^\alpha]\,\hat{\rho}_0\big)E^\alpha(t) \tag{D.5}$$

which translates into a conductivity

$$\begin{aligned}\sigma^{\beta\alpha}(\bar{\omega}) &= \frac{i\,e}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}^\alpha(\bar{\omega})\Big) + \frac{i\,e^2}{\hbar}\,\mathrm{Tr}\big([\hat{r}^\beta, \hat{r}^\alpha]\,\hat{\rho}_0\big) \\ &= -\frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\big[D^\beta, H_0\big]_{\mathbf{k}ba}}{\hbar\omega - \Delta\epsilon_{\mathbf{k}ab}}\,[D^\alpha, \rho_0]_{\mathbf{k}ab} - \frac{i\,e^2}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a [D^\beta, D^\alpha]_{\mathbf{k}aa}\,f_{\mathbf{k}a}\end{aligned} \tag{D.6}$$

and, in general,

$$J^{\beta(n)}(t) = \frac{i\,e}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta, \hat{H}_0]\,\hat{\rho}^{(n)}(t)\Big) + \frac{i\,e^2}{\hbar}\,\mathrm{Tr}\big([\hat{r}^\beta, \hat{r}^\alpha]\,\hat{\rho}^{(n-1)}(t)\big)E^\alpha(t) \tag{D.7}$$

which implies,

$$\sigma^{\beta\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)$$
$$= \frac{i\,e}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta,\hat{H}_0]\,\hat{\rho}^{\alpha_1\ldots\alpha_n}(\bar{\omega}_1,\ldots,\bar{\omega}_n)\Big) + \frac{i\,e^2}{\hbar}\,\mathrm{Tr}\Big([\hat{r}^\beta,\hat{r}^{\alpha_n}]\,\hat{\rho}^{\alpha_1\ldots\alpha_{n-1}}(\bar{\omega}_1,\ldots,\bar{\omega}_{n-1})\Big) \tag{D.8}$$

with the density matrix components given by

$$\rho^{\alpha_1\ldots\alpha_n}_{\mathbf{k}ab}(\bar{\omega}_1,\ldots,\bar{\omega}_n) = \frac{(i\,e)^n}{\hbar\bar{\omega}_1+\cdots+\hbar\bar{\omega}_n-\Delta\epsilon}\circ\left[D^{\alpha_n},\ldots\frac{1}{\hbar\bar{\omega}_1-\Delta\epsilon}\circ[D^{\alpha_1},\rho_0]\ldots\right]_{\mathbf{k}ab} \tag{D.9}$$

same as before.

In summary, the perturbation theory of the density matrix itself remained unchanged, the only alteration being in the definition of the current operator. Due to the noncommutative covariant derivatives, there is an extra term in the nonlinear conductivity, with a complexity comparable to the previous order nonlinear conductivity.

This is in line with the discussion from Appendix B: Eq. D.4 is a special case of Eq. B.21. The nonlinear conductivity that follows from Eq. D.8 may be more complicated than its counterpart in Section 5.2, but is equally amenable to a resonance-based decomposition, presented in Appendix E.

Appendix E

Resonance-based analysis without commuting position operators

In this appendix, the results of Section 5.2 are revised and set at a slightly more general level, by dropping the assumption of commutative covariant derivatives (Appendices C and D). This has the advantage of making certain properties, such as overall permutation symmetry (Chapter 6) easier to check.

E.1 Linear order

The optical conductivity has the following resonance-based decomposition,

$$\sigma^{\beta\alpha_1}(\bar{\omega}_1) = \sigma_F^{\beta\alpha_1}(\bar{\omega}_1) + \sigma_1^{\beta\alpha_1}(\bar{\omega}_1) \tag{E.1}$$

with

$$\sigma_F^{\beta\alpha_1}(\bar{\omega}_1) = \frac{i\,e^2}{\hbar}\left(-\frac{1}{\hbar\bar{\omega}_1}\,F_A^{\beta\alpha_1} + i\,F_B^{\beta\alpha_1}\right) \tag{E.2}$$

$$\sigma_1^{\beta\alpha_1}(\bar{\omega}_1) = \frac{i\,e^2}{\hbar}\,\hbar\bar{\omega}_1\,\Pi_1^{\beta\alpha_1}(\bar{\omega}_1) \tag{E.3}$$

At linear order, there are no differences when compared with Section 5.2.1.

E.2 Second order

The second order optical conductivity has the following resonance-based decomposition,

$$\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \sigma_F^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) + \sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) + \sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) \tag{E.4}$$

into the Fermi surface contribution,

$$\sigma_F^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \frac{i\,e^3}{\hbar}\left(\frac{i}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2}\,F_A^{\beta\alpha_1\alpha_2} - \frac{1}{\hbar\bar{\omega}_2}\,F_B^{\alpha_1\alpha_2\beta}\right) \tag{E.5}$$

the one-photon contribution,

$$\frac{\hbar}{i\,e^3}\,\sigma_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}\,\Pi_1^{\alpha_2\alpha_1\beta}(\bar{\omega}_1) + \frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1) + \frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\,\Pi_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1) + \Pi_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1) - \Pi_2^{\alpha_1\alpha_2\beta}(-\bar{\omega}_1) \tag{E.6}$$

and the two-photon contribution,

$$\frac{\hbar}{i\,e^3}\,\sigma_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = -\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) + \frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) \tag{E.7}$$

E.3 Third order

The third order optical conductivity has the following resonance-based decomposition,

$$\sigma^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \sigma_F^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) + \sigma_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) + \sigma_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) + \sigma_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) \tag{E.8}$$

Once again, all the contributions will be written in detail.

The Fermi surface contribution,

$$\sigma_F^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \frac{i\,e^4}{\hbar}\left(\frac{1}{6\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3}\,F_A^{\beta\alpha_1\alpha_2\alpha_3} + \frac{i}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}\,F_B^{\alpha_2\beta\alpha_1\alpha_3}\right) \tag{E.9}$$

The one-photon contribution,

$$\begin{aligned}
&\frac{\hbar}{i\,e^4}\,\sigma_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) = \\
&+\frac{1}{\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{123}}\,\Pi_1^{\alpha_1\alpha_2\alpha_3\beta}(-\bar{\omega}_1) + \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_1) - \frac{1}{(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\alpha_1\alpha_2\beta\alpha_3}(-\bar{\omega}_1) \\
&-\frac{1}{\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_3}\,\Pi_2^{\alpha_2\alpha_1\alpha_3\beta}(\bar{\omega}_1) - \frac{\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1) - \frac{1}{(\hbar\bar{\omega}_3)^2\,\hbar\bar{\omega}_{123}}\,\Pi_3^{\alpha_2\alpha_1\alpha_3\beta}(\bar{\omega}_1) \\
&-\frac{\hbar\bar{\omega}_{123}}{2\,(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1) + \frac{\hbar\bar{\omega}_1(\hbar\bar{\omega}_2-\hbar\bar{\omega}_3)}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1) - \frac{1}{\hbar\bar{\omega}_3}\,\Pi_4^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_1) \\
&-\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_{23}}\,\Pi_4^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}_1) + \frac{1}{\hbar\bar{\omega}_3}\,\Pi_4^{\alpha_1\alpha_2\beta\alpha_3}(-\bar{\omega}_1) - \frac{1}{\hbar\bar{\omega}_{12}}\,\Pi_4^{\alpha_2\alpha_1\beta\alpha_3}(\bar{\omega}_1) \\
&+\frac{1}{\hbar\bar{\omega}_{12}}\,\Pi_4^{\alpha_2\alpha_1\alpha_3\beta}(\bar{\omega}_1) + \frac{2\,\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_{13}\,\hbar\bar{\omega}_{23}}\,\Pi_5^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_1) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_{12}\,\hbar\bar{\omega}_{13}}\,\Pi_5^{\alpha_2\alpha_3\alpha_1\beta}(\bar{\omega}_1)
\end{aligned} \tag{E.10}$$

with the abbreviations $\bar{\omega}_{ij} \equiv \bar{\omega}_i + \bar{\omega}_j$ and $\bar{\omega}_{123} \equiv \bar{\omega}_1 + \bar{\omega}_2 + \bar{\omega}_3$.

The two-photon contribution,

$$\frac{\hbar}{i\,e^4}\,\sigma_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) =$$
$$+\,\frac{1}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\alpha_1\alpha_3\beta\alpha_2}(-\bar{\omega}_{12}) - \frac{\hbar\bar{\omega}_{123}}{(\hbar\bar{\omega}_2)^2\,\hbar\bar{\omega}_3}\,\Pi_1^{\alpha_1\beta\alpha_3\alpha_2}(-\bar{\omega}_{12}) - \frac{1}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_{123}}\,\Pi_1^{\alpha_3\alpha_1\alpha_2\beta}(\bar{\omega}_{12})$$
$$-\,\frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{12}) + \frac{1}{(\hbar\bar{\omega}_1)^2\,\hbar\bar{\omega}_{123}}\Pi_3^{\alpha_3\alpha_2\alpha_1\beta}(\bar{\omega}_{12}) + \frac{\hbar\bar{\omega}_{123}}{(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{12})$$
$$+\,\frac{1}{\hbar\bar{\omega}_2}\,\Pi_6^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_{12}) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3}\,\Pi_6^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}_{12}) \tag{E.11}$$

The three-photon contribution,

$$\frac{\hbar}{i\,e^4}\,\sigma_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_1,\bar{\omega}_2,\bar{\omega}_3) =$$
$$+\,\frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_2(\hbar\bar{\omega}_3)^2}\,\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) + \frac{\hbar\bar{\omega}_{123}}{2\,\hbar\bar{\omega}_2\,\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123})$$
$$-\,\frac{\hbar\bar{\omega}_{123}}{2\,(\hbar\bar{\omega}_2)^2\,(\hbar\bar{\omega}_3)^2}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_3\,\hbar\bar{\omega}_{23}}\,\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}_{123}) - \frac{\hbar\bar{\omega}_{123}}{\hbar\bar{\omega}_{13}\,\hbar\bar{\omega}_{23}}\,\Pi_5^{\alpha_3\beta\alpha_1\alpha_2}(\bar{\omega}_{123}) \tag{E.12}$$

Appendix F

Resonance-based decomposition of the second order conductivity

As an illustration of the scheme by which the expressions in Sections 5.2.1-5.2.3 were derived, the case of the second order conductivity is treated here in detail. The aim is to arrive at the results of Section 5.2.2, starting from the general expression for a nonlinear conductivity. The derivation of the third order conductivity, albeit considerably lengthier, follows along the same lines.

We start with Eq. 5.3, taken for the case $n = 2$,

$$\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1, \bar{\omega}_2) = \frac{e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\left[D^\beta, H_0\right]_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab}} \left[D^{\alpha_2}, \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right]_{\mathbf{k}ab} \tag{F.1}$$

and separate the diagonal and off-diagonal current matrix elements,

$$\begin{aligned}\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1, \bar{\omega}_2) = {} & \frac{e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a} \frac{(\partial^\beta \epsilon_{\mathbf{k}a})}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2} \left[D^{\alpha_2}, \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right]_{\mathbf{k}aa} \\ & - \frac{i\,e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a\neq b} \frac{\mathcal{A}^\beta_{\mathbf{k}ba}\,\Delta\epsilon_{\mathbf{k}ab}}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab}} \left[D^{\alpha_2}, \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right]_{\mathbf{k}ab}\end{aligned} \tag{F.2}$$

Eq. F.2 is expressed in a condensed notation. To expand it into more explicit formulae, Eq. 2.52 is applied iteratively, starting with

$$[D^\alpha, \rho_0]_{\mathbf{k}ab} = \delta_{ab}\,(\partial^\alpha f_{\mathbf{k}a}) - i\,\mathcal{A}^\alpha_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba} \tag{F.3}$$

where δ_{ab} is the Kronecker delta and $f_{\mathbf{k}a} = f(\epsilon_{\mathbf{k}a})$ is the Fermi function.

In the intermediate steps, the definition of the Hadamard product is used. For instance,

$$\left(\frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right)_{\mathbf{k}aa} = \frac{[D^{\alpha_1}, \rho_0]_{\mathbf{k}aa}}{\hbar\bar{\omega}_1} = \frac{(\partial^{\alpha_1} f_{\mathbf{k}a})}{\hbar\bar{\omega}_1} \tag{F.4}$$

$$\left(\frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right)_{\mathbf{k}ab} = \frac{[D^{\alpha_1}, \rho_0]_{\mathbf{k}ab}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} = -\frac{i\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} \tag{F.5}$$

with $a \neq b$ in the last equation.

Expanding the first commutator in Eq. F.2,

$$\left[D^{\alpha_2}, \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right]_{\mathbf{k}aa} =$$

$$\frac{\partial^{\alpha_2}\partial^{\alpha_1} f_{\mathbf{k}a}}{\hbar\,\bar{\omega}_1} + \sum_b \left(\frac{\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ba}\,\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 + \Delta\epsilon_{\mathbf{k}ab}} + \frac{\mathcal{A}^{\alpha_2}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} \right) \tag{F.6}$$

while the second commutator gives, for $a \neq b$,

$$\left[D^{\alpha_2}, \frac{1}{\hbar\bar{\omega}_1 - \Delta\epsilon} \circ [D^{\alpha_1}, \rho_0]\right]_{\mathbf{k}ab} = -\frac{i}{\hbar\bar{\omega}_1}\,\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\,(\partial^{\alpha_1}\,\Delta f_{\mathbf{k}ba})$$

$$-\frac{i\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab;\alpha_2}\,\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} - \frac{i\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,(\partial^{\alpha_2}\Delta f_{\mathbf{k}ba})}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} - \frac{i\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab})\,\Delta f_{\mathbf{k}ba}}{(\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab})^2} \tag{F.7}$$

Replacing Eqs. F.6 and F.7 in Eq. F.2,

$$\begin{aligned}
\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1, \bar{\omega}_2) = {} & \frac{e^3}{\hbar}\,\frac{1}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2)} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_a \left(\partial^\beta \epsilon_{\mathbf{k}a}\right) \left(\partial^{\alpha_1}\partial^{\alpha_2} f_{\mathbf{k}a}\right) \\
& + \frac{e^3}{\hbar}\,\frac{1}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\alpha_2}_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}(\partial^\beta\Delta\epsilon_{\mathbf{k}ab})\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab}} \\
& - \frac{e^3}{\hbar}\,\frac{1}{\hbar\bar{\omega}_1} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\Delta\epsilon_{\mathbf{k}ab}(\partial^{\alpha_1}\Delta f_{\mathbf{k}ba})}{\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab}} \\
& - \frac{e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\Delta\epsilon_{\mathbf{k}ab}(\partial^{\alpha_2}\Delta f_{\mathbf{k}ba})}{(\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab})(\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab})} \\
& - \frac{e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab;\alpha_2}\,\Delta f_{\mathbf{k}ba}}{(\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab})(\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab})} \\
& - \frac{e^3}{\hbar} \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,(\partial^{\alpha_2}\Delta\epsilon_{ab})\,\Delta\epsilon_{\mathbf{k}ab}\,\Delta f_{\mathbf{k}ba}}{(\hbar\bar{\omega}_1 - \Delta\epsilon_{\mathbf{k}ab})^2(\hbar\bar{\omega}_1 + \hbar\bar{\omega}_2 - \Delta\epsilon_{\mathbf{k}ab})}
\end{aligned} \tag{F.8}$$

The next step is a partial fraction decomposition,

$$
\begin{aligned}
&\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \\
&+\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1\left(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2\right)}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_a\left(\partial^\beta\epsilon_{\mathbf{k}a}\right)\left(\partial^{\alpha_1}\partial^{\alpha_2}f_{\mathbf{k}a}\right)\\
&+\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\frac{\mathcal{A}^{\alpha_2}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\left(\partial^\beta\Delta\epsilon_{\mathbf{k}ab}\right)\Delta f_{\mathbf{k}ba}}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}\\
&-\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\left(\partial^{\alpha_1}\Delta f_{\mathbf{k}ba}\right)\left(\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2-\Delta\epsilon_{\mathbf{k}ab}}-1\right)\\
&-\frac{e^3}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\left(\partial^{\alpha_2}\Delta f_{\mathbf{k}ba}\right)\\
&\qquad\qquad\times\left(\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\frac{1}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2-\Delta\epsilon_{\mathbf{k}ab}}\right)\\
&-\frac{e^3}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab;\alpha_2}\,\Delta f_{\mathbf{k}ba}\\
&\qquad\qquad\times\left(\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\frac{1}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2-\Delta\epsilon_{\mathbf{k}ab}}\right)\\
&-\frac{e^3}{\hbar}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\left(\partial^{\alpha_2}\Delta\epsilon_{ab}\right)\Delta f_{\mathbf{k}ba}\\
&\times\left(\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\frac{1}{(\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab})^2}-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\frac{1}{\hbar\bar{\omega}_1-\Delta\epsilon_{\mathbf{k}ab}}+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2-\Delta\epsilon_{\mathbf{k}ab}}\right)
\end{aligned}
\tag{F.9}
$$

The second term in the third integral (fourth line) of Eq. F.9 can be singled out and identified as a Fermi surface integral,

$$
\begin{aligned}
&\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\left(\partial^{\alpha_1}\Delta f_{\mathbf{k}ba}\right)\\
&=\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a,b}\left(\mathcal{A}^{\beta}_{\mathbf{k}ab}\mathcal{A}^{\alpha_2}_{\mathbf{k}ba}-\mathcal{A}^{\beta}_{\mathbf{k}ba}\mathcal{A}^{\alpha_2}_{\mathbf{k}ab}\right)\left(\partial^{\alpha_1}f_{\mathbf{k}a}\right)\\
&=\frac{e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a}\left[\mathcal{A}^{\beta},\mathcal{A}^{\alpha_2}\right]_{\mathbf{k}aa}\left(\partial^{\alpha_1}f_{\mathbf{k}a}\right)\\
&=-\frac{i\,e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}\int\frac{d^d\mathbf{k}}{(2\pi)^d}\sum_{a}\mathcal{F}^{\beta\alpha_2}_{a}\left(\partial^{\alpha_1}f_{\mathbf{k}a}\right)\\
&=-\frac{i\,e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}F_B^{\alpha_1\beta\alpha_2}\\
&=+\frac{i\,e^3}{\hbar}\frac{1}{\hbar\bar{\omega}_1}F_B^{\alpha_1\alpha_2\beta}
\end{aligned}
\tag{F.10}
$$

with the use of Eq. C.6.

Identifying the remaining integrals with the definitions in Eqs. 5.26-5.29 and G.2,

$$
\begin{aligned}
&\frac{\hbar}{i\,e^3}\,\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \\
&+\frac{i}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}\,F_A^{\beta\alpha_1\alpha_2}+\frac{1}{\hbar\bar{\omega}_1}\,F_B^{\alpha_1\alpha_2\beta}-\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\,\Pi_B^{\beta\alpha_1\alpha_2}(\bar{\omega}_1) \\
&-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\,\Pi_B^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2)+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_1}\,\Pi_B^{\beta\alpha_2\alpha_1}(\bar{\omega}_1+\bar{\omega}_2)+\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}\,\Pi_1^{\alpha_2\alpha_1\beta}(\bar{\omega}_1) \\
&+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1)+\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\left(\frac{d}{d\bar{\omega}_1}\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1)\right)-\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1) \\
&-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2)+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) \qquad \text{(F.11)}
\end{aligned}
$$

The first term is symmetrical in all tensor indices and can be rewritten as

$$
\begin{aligned}
\frac{i}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}\,F_A^{\beta\alpha_1\alpha_2} &\to \frac{1}{2}\left(\frac{i}{\hbar\bar{\omega}_1\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}+\frac{i}{\hbar\bar{\omega}_2\,(\hbar\bar{\omega}_1+\hbar\bar{\omega}_2)}\right)F_A^{\beta\alpha_1\alpha_2} \\
&= \frac{i}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2}\,F_A^{\beta\alpha_1\alpha_2} \qquad \text{(F.12)}
\end{aligned}
$$

by use of intrinsic permutation symmetry. The same symmetry is responsible for cancelling the terms in $\Pi_B(\bar{\omega}_1+\bar{\omega}_2)$ in Eq. F.11.

Applying Eq. G.1 to replace the derivative in Eq. F.11, we obtain

$$
\begin{aligned}
&\frac{\hbar}{i\,e^3}\,\sigma^{\beta\alpha_1\alpha_2}(\bar{\omega}_1,\bar{\omega}_2) = \\
&\quad+\frac{i}{2\,\hbar\bar{\omega}_1\,\hbar\bar{\omega}_2}\,F_A^{\beta\alpha_1\alpha_2}+\frac{1}{\hbar\bar{\omega}_1}\,F_B^{\alpha_1\alpha_2\beta} \\
&\quad+\frac{1}{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}\,\Pi_1^{\alpha_2\alpha_1\beta}(\bar{\omega}_1)+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1)+\frac{\hbar\bar{\omega}_1}{\hbar\bar{\omega}_2}\,\Pi_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}_1) \\
&\quad-\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{(\hbar\bar{\omega}_2)^2}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2)+\frac{\hbar\bar{\omega}_1+\hbar\bar{\omega}_2}{\hbar\bar{\omega}_2}\,\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}_1+\bar{\omega}_2) \qquad \text{(F.13)}
\end{aligned}
$$

in agreement with Eqs. 5.22-5.29.

Appendix G

Integral identities

In the derivation of the expressions in Sections 5.2.1-5.2.3, integral identities were used that follow directly from the definitions in Eqs. 5.19-5.21, 5.26-5.29, 5.35-5.42. This is exemplified by the decomposition of the second order conductivity in Appendix F. Additional identities emerge under specific assumptions such as commutative covariant derivatives (Appendix C) or time-reversal symmetry (Section 6.6). All these integral identities are helpful in manipulating expressions and are summarized here.

The first set of identities concerns derivatives of Π_j^α integrals. In second order,

$$\frac{d}{d\bar{\omega}}\,\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}) = \Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}) + \Pi_2^{\alpha_1\beta\alpha_2}(-\bar{\omega}) + \Pi_B^{\beta\alpha_1\alpha_2}(\bar{\omega}) \tag{G.1}$$

with

$$i\,\Pi_B^{\beta\alpha_1\alpha_2}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,(\partial^{\alpha_2}\Delta f_{\mathbf{k}ba})}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \tag{G.2}$$

which vanishes in the absence of a Fermi surface.

In third order,

$$\frac{d}{d\bar{\omega}}\,\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) + \Pi_6^{\beta\alpha_3\alpha_1\alpha_2}(\bar{\omega}) + \Pi_B^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.3}$$

$$\begin{aligned}\frac{d}{d\bar{\omega}}\,\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) &= \Pi_1^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) - \Pi_1^{\alpha_1\beta\alpha_3\alpha_2}(-\bar{\omega}) + \Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) + \Pi_C^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \\ &= \Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) - \Pi_1^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega}) + \Pi_2^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) + \Pi_C^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega})\end{aligned} \tag{G.4}$$

with

$$\Pi_B^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab;\alpha_2}\,(\partial^{\alpha_3}\Delta f_{\mathbf{k}ba})}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \tag{G.5}$$

$$\Pi_C^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \equiv \int \frac{d^d\mathbf{k}}{(2\pi)^d} \sum_{a,b} \frac{\mathcal{A}^{\beta}_{\mathbf{k}ba}\,\mathcal{A}^{\alpha_1}_{\mathbf{k}ab}\,(\partial^{\alpha_2}\Delta\epsilon_{\mathbf{k}ab})\,(\partial^{\alpha_3}\Delta f_{\mathbf{k}ba})}{\hbar\bar{\omega} - \Delta\epsilon_{\mathbf{k}ab}} \tag{G.6}$$

which vanish in the absence of a Fermi surface.

Additionally, it follows from Eq. G.4,

$$\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega})-\Pi_1^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega})-\Pi_1^{\alpha_1\beta\alpha_2\alpha_3}(-\bar{\omega})+\Pi_1^{\alpha_1\beta\alpha_3\alpha_2}(-\bar{\omega}) = \Pi_C^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega})-\Pi_C^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) \tag{G.7}$$

The second set of identities, unlike the first, is reliant on time-reversal symmetry and consists on simply stating the parity of the integrals in Eqs. 5.21, 5.28-5.29, 5.37-5.42,

$$\Pi_1^{\beta\alpha_1}(-\bar{\omega}) = \Pi_1^{\beta\alpha_1}(\bar{\omega}) \tag{G.8}$$
$$\Pi_1^{\beta\alpha_1\alpha_2}(-\bar{\omega}) = \Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}) \tag{G.9}$$
$$\Pi_2^{\beta\alpha_1\alpha_2}(-\bar{\omega}) = -\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}) \tag{G.10}$$
$$\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = -\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.11}$$
$$\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = -\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.12}$$
$$\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = \Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.13}$$
$$\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = \Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.14}$$
$$\Pi_5^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = \Pi_5^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.15}$$
$$\Pi_6^{\beta\alpha_1\alpha_2\alpha_3}(-\bar{\omega}) = \Pi_6^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) \tag{G.16}$$

Finally, there is a third set of identities, relating different tensor elements of an integral,

$$F_A^{\beta\alpha_1} = F_A^{\alpha_1\beta} \tag{G.17}$$
$$F_B^{\beta\alpha_1} = -F_B^{\alpha_1\beta} \tag{G.18}$$
$$\Pi_1^{\beta\alpha_1}(\bar{\omega}) = \Pi_1^{\alpha_1\beta}(\bar{\omega}) \tag{G.19}$$
$$F_A^{\beta\alpha_1\alpha_2} = F_A^{\alpha_1\beta\alpha_2} = F_A^{\alpha_1\alpha_2\beta} \tag{G.20}$$
$$F_B^{\beta\alpha_1\alpha_2} = -F_B^{\beta\alpha_2\alpha_1} \tag{G.21}$$
$$\Pi_1^{\beta\alpha_1\alpha_2}(\bar{\omega}) = -\Pi_1^{\alpha_1\beta\alpha_2}(\bar{\omega}) \tag{G.22}$$
$$\Pi_2^{\beta\alpha_1\alpha_2}(\bar{\omega}) = \Pi_2^{\beta\alpha_2\alpha_1}(\bar{\omega}) \tag{G.23}$$
$$F_A^{\beta\alpha_1\alpha_2\alpha_3} = F_A^{\alpha_1\beta\alpha_2\alpha_3} = F_A^{\alpha_1\alpha_2\beta\alpha_3} = F_A^{\alpha_1\alpha_2\alpha_3\beta} \tag{G.24}$$
$$F_B^{\beta\alpha_1\alpha_2\alpha_3} = F_B^{\alpha_3\alpha_1\alpha_2\beta} = -F_B^{\beta\alpha_2\alpha_1\alpha_3} = -F_B^{\alpha_3\alpha_2\alpha_1\beta} \tag{G.25}$$
$$\Pi_1^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_1^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}) \tag{G.26}$$
$$\Pi_2^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_2^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) = \Pi_2^{\alpha_1\beta\alpha_2\alpha_3}(\bar{\omega}) \tag{G.27}$$
$$\Pi_3^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_3^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) = \Pi_3^{\alpha_1\beta\alpha_2\alpha_3}(\bar{\omega}) \tag{G.28}$$
$$\Pi_4^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_4^{\beta\alpha_2\alpha_1\alpha_3}(\bar{\omega}) \tag{G.29}$$
$$\Pi_5^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_5^{\alpha_1\beta\alpha_2\alpha_3}(\bar{\omega}) = \Pi_5^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) = \Pi_5^{\alpha_2\alpha_3\beta\alpha_1}(\bar{\omega}) \tag{G.30}$$
$$\Pi_6^{\beta\alpha_1\alpha_2\alpha_3}(\bar{\omega}) = \Pi_6^{\alpha_1\beta\alpha_2\alpha_3}(\bar{\omega}) = \Pi_6^{\beta\alpha_1\alpha_3\alpha_2}(\bar{\omega}) = \Pi_6^{\alpha_2\alpha_3\beta\alpha_1}(\bar{\omega}) \tag{G.31}$$

These can be seen by inspection of the integrands. Eqs. G.23, G.26, G.29 and G.31 are valid only when commutativity of covariant derivatives is assumed (Appendix C), since they are a direct consequence of Eq. C.7. Eqs. G.19 and G.22, as well as the last equalities in Eqs. G.27, G.28, G.30, G.31, are reliant on time-reversal symmetry.

Appendix H

Integral evaluation for monolayer graphene

The evaluation of the integrals in Eqs. 5.19-5.21, 5.26-5.29 and 5.35-5.42 for monolayer graphene near the Dirac point is relatively straightforward. Eqs 3.73 for the band energies and Eqs. 3.75-3.78 for the matrix elements of the non-Abelian Berry connection are substituted and their derivatives (or generalized derivatives) are computed.

Here, the possibility of a band gap ($\Delta \neq 0$) is considered, with the effective gap defined as usual: $\Delta_{eff} \equiv \max(2|\mu|, \Delta)$. Taking the relaxation-free limit and setting $T = 0$, the integration in $\mathcal{I}_j^{\alpha}$ can be done easily, by converting the integral over the FBZ to polar coordinates and using the Dirac delta function in Eq. 5.45.

For example, in linear order,

$$\begin{aligned}
\mathcal{I}_1^{xx}(\omega) &= 4 \int \frac{d^2\mathbf{q}}{(2\pi)^2} \sum_{a,b} \mathcal{A}^x_{\mathbf{q}ba}\, \mathcal{A}^x_{\mathbf{q}ab}\, \Delta f_{\mathbf{q}ba}\, \delta\left(\hbar\omega - \Delta\epsilon_{\mathbf{q}ab}\right) \\
&= 4 \int \frac{d^2\mathbf{q}}{(2\pi)^2}\, \mathcal{A}^x_{\mathbf{q}cv}\, \mathcal{A}^x_{\mathbf{q}vc}\, \Delta f_{\mathbf{q}vc} \left(\delta\left(\hbar\omega - \Delta\epsilon_{\mathbf{q}cv}\right) - \delta\left(\hbar\omega + \Delta\epsilon_{\mathbf{q}cv}\right)\right) \\
&= 4 \int \frac{d^2\mathbf{q}}{(2\pi)^2}\, |\mathcal{A}^x_{\mathbf{q}cv}|^2\, \mathrm{sign}(\omega)\, \delta\left(|\hbar\omega| - \Delta\epsilon_{\mathbf{q}cv}\right)\, \Theta\left(\Delta\epsilon_{\mathbf{q}cv} - \Delta_{eff}\right)
\end{aligned} \tag{H.1}$$

The factor of 4 in the beginning takes into account the valley and spin degeneracies in the system. The Heaviside step function at the end comes from considering the difference in Fermi functions $\Delta f_{\mathbf{q}vc}$ at $T = 0$, as well as the absence of states below the band gap. The two Dirac deltas were joined, by noting that the first Dirac delta is only nonzero for positive frequencies and the second for negative ones. The validity of the manipulation can be checked by examining the cases $\omega > 0$ and $\omega < 0$, separately.

Replacing the non-Abelian Berry connection by the matrix element in Eq. 3.77 and inserting the dispersion relation in Eq. 3.73,

$$\begin{aligned}
\mathcal{I}_1^{xx}(\omega) = 4 \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^{+\infty} \frac{dq}{2\pi}\, q\, \frac{\sin^2\theta\ \mathrm{sign}(\omega)}{4\left(q^2 + (\Delta/2\hbar v_F)^2\right)}\, \delta\left(|\hbar\omega| - \sqrt{(2v_F\hbar\, q)^2 + \Delta^2}\right) \Theta\left(|\hbar\omega| - \Delta_{eff}\right) \\
+ 4 \int_0^{2\pi} \frac{d\theta}{2\pi} \int_0^{+\infty} \frac{dq}{2\pi}\, q\, \frac{(\Delta/2\hbar v_F)^2\ \cos^2\theta\ \mathrm{sign}(\omega)}{4\left(q^2 + (\Delta/2\hbar v_F)^2\right)^2}\, \delta\left(|\hbar\omega| - \sqrt{(2v_F\hbar\, q)^2 + \Delta^2}\right) \Theta\left(|\hbar\omega| - \Delta_{eff}\right)
\end{aligned} \tag{H.2}$$

where advantage was taken of the isotropy of the dispersion relation, by writing the integration in polar coordinates $\mathbf{q} = (q_x, q_y) = q\,(\cos\theta, \sin\theta)$.

With the variable change $q' \equiv (2\hbar v_F)\, q$,

$$\begin{aligned}
&\mathcal{I}_1^{xx}(\omega) = \\
&+ 4\int_0^{2\pi} \frac{d\theta}{2\pi}\, \sin^2\theta \int_0^{+\infty} \frac{dq'}{2\pi}\, \frac{q'}{4\,(q'^2+\Delta^2)}\, \mathrm{sign}(\omega)\, \delta\left(|\hbar\omega| - \sqrt{q'^2+\Delta^2}\right)\, \Theta\left(|\hbar\omega| - \Delta_{eff}\right) \\
&+ 4\int_0^{2\pi} \frac{d\theta}{2\pi}\, \cos^2\theta \int_0^{+\infty} \frac{dq'}{2\pi}\, \frac{q'\,\Delta^2}{4\,(q'^2+\Delta^2)^2}\, \mathrm{sign}(\omega)\, \delta\left(|\hbar\omega| - \sqrt{q'^2+\Delta^2}\right)\, \Theta\left(|\hbar\omega| - \Delta_{eff}\right)
\end{aligned} \tag{H.3}$$

At this point, the identity $\delta(g(x)) = \delta(x-a)/|g'(a)|$ for $g(a) = 0$, where a is the only zero of the function $g(x)$, is used to simplify the Dirac delta function,

$$\begin{aligned}
\mathcal{I}_1^{xx}(\omega) = {} & 4\left(\frac{1}{2}\right)\int_0^{+\infty} \frac{dq'}{2\pi}\, \frac{q'}{4\,(q'^2+\Delta^2)}\, \mathrm{sign}(\omega)\, \frac{\delta\left(q' - \sqrt{(\hbar\omega)^2-\Delta^2}\right)}{\sqrt{(\hbar\omega)^2-\Delta^2}/|\hbar\omega|}\, \Theta\left(|\hbar\omega| - \Delta_{eff}\right) \\
& + 4\left(\frac{1}{2}\right)\int_0^{+\infty} \frac{dq'}{2\pi}\, \frac{q'\,\Delta^2}{4\,(q'^2+\Delta^2)^2}\, \mathrm{sign}(\omega)\, \frac{\delta\left(q' - \sqrt{(\hbar\omega)^2-\Delta^2}\right)}{\sqrt{(\hbar\omega)^2-\Delta^2}/|\hbar\omega|}\, \Theta\left(|\hbar\omega| - \Delta_{eff}\right)
\end{aligned} \tag{H.4}$$

Now, the integration with the Dirac delta leads to direct substitutions. The $\mathcal{I}_1^{xx}(\omega)$ integral is then provided by the following analytical expression,

$$\mathcal{I}_1^{xx}(\omega) = \frac{1}{\pi}\left(\frac{\hbar\omega}{4(\hbar\omega)^2} + \frac{\hbar\omega\,\Delta^2}{4\,(\hbar\omega)^4}\right)\Theta\left(|\hbar\omega| - \Delta_{eff}\right) = \frac{1}{4\pi\hbar\omega}\left(1 + \left(\frac{\Delta}{\hbar\omega}\right)^2\right)\Theta(\hbar|\omega| - \Delta_{eff}) \tag{H.5}$$

whose Hilbert transform gives

$$\mathcal{H}_1^{xx}(\omega) = \frac{1}{4\pi^2\hbar\omega}\left(1 + \left(\frac{\Delta}{\hbar\omega}\right)^2\right)\ln\left|\frac{\hbar\omega - \Delta_{eff}}{\hbar\omega + \Delta_{eff}}\right| + \frac{\Delta^2}{2\pi^2\,(\hbar\omega)^2\,\Delta_{eff}} \tag{H.6}$$

From the crystal symmetry, it is known that

$$\mathcal{I}_1^{yy}(\omega) = \mathcal{I}_1^{xx}(\omega) \tag{H.7}$$
$$\mathcal{H}_1^{yy}(\omega) = \mathcal{H}_1^{xx}(\omega) \tag{H.8}$$
$$\mathcal{I}_1^{xy}(\omega) = \mathcal{I}_1^{yx}(\omega) = 0 \tag{H.9}$$
$$\mathcal{H}_1^{xy}(\omega) = \mathcal{H}_1^{yx}(\omega) = 0 \tag{H.10}$$

as can be confirmed by performing the integrations for the other tensor elements.

There are also the F_X^α integrals, required to describe the Fermi surface contribution,

$$\begin{aligned} F_A^{xx} &= +4 \int \frac{d^2\mathbf{q}}{(2\pi)^2} \sum_a \left(\partial^x \epsilon_{\mathbf{q}a}\right) \left(\partial^x f_{\mathbf{q}a}\right) \\ &= -4 \int \frac{d^2\mathbf{q}}{(2\pi)^2} \sum_a \left(\partial^x \epsilon_{\mathbf{q}a}\right)^2 \delta(\mu - \epsilon_{\mathbf{q}a}) \\ &= -4 \int \frac{d^2\mathbf{q}}{(2\pi)^2} \left(\partial^x \epsilon_{\mathbf{q}c}\right)^2 \delta(|\mu| - \epsilon_{\mathbf{q}c}) \end{aligned} \tag{H.11}$$

where the Fermi function derivative was assessed for $T = 0$ and, in the last step, the electron-hole symmetry $\epsilon_{\mathbf{q}v} = -\epsilon_{\mathbf{q}c}$ was used.

Replacing with Eq. 3.73,

$$\begin{aligned} F_A^{xx} &= -4 \int_0^{2\pi} \frac{d\theta}{(2\pi)} \int_0^{+\infty} \frac{dq}{(2\pi)}\, q \left(\frac{1}{2} \frac{(2\hbar v_F)^2\, q\, \cos\theta}{\sqrt{(2\hbar v_F\, q)^2 + \Delta^2}} \right)^2 \delta(|\mu| - \sqrt{(2\hbar v_F\, q)^2 + \Delta^2}/2) \\ &= -4 \int_0^{2\pi} \frac{d\theta}{(2\pi)} \cos^2\theta \int_0^{+\infty} \frac{dq'}{(2\pi)}\, q' \left(\frac{1}{2} \frac{q'}{\sqrt{q'^2 + \Delta^2}} \right)^2 \delta(|\mu| - \sqrt{q'^2 + \Delta^2}/2) \end{aligned} \tag{H.12}$$

The Dirac delta is once more simplified by finding the zero of its argument. Notice that, in this case, a solution exists only if twice the modulus of the chemical potential exceeds the band gap (in other words, if the Fermi level crosses either the valence or the conduction band), leading to the appearance of an Heaviside step function,

$$\begin{aligned} F_A^{xx} &= -4 \left(\frac{1}{2}\right) \int_0^{+\infty} \frac{dq'}{(2\pi)}\, q' \left(\frac{1}{2} \frac{q'}{\sqrt{q'^2 + \Delta^2}} \right)^2 \frac{\delta(q' - \sqrt{(2\mu)^2 - \Delta^2})}{\sqrt{(2\mu)^2 - \Delta^2}/4|\mu|}\, \Theta(2|\mu| - \Delta) \\ &= -2\, \frac{1}{2\pi} \sqrt{(2\mu)^2 - \Delta^2} \left(\frac{\sqrt{(2\mu)^2 - \Delta^2}}{4|\mu|} \right)^2 \frac{4|\mu|}{\sqrt{(2\mu)^2 - \Delta^2}}\, \Theta(2|\mu| - \Delta) \\ &= -\frac{(2\mu)^2 - \Delta^2}{4\pi\, |\mu|}\, \Theta(2|\mu| - \Delta) \\ &= -\frac{|\mu|}{\pi} \left(1 - \left(\frac{\Delta}{2\,\mu} \right)^2 \right) \Theta(2|\mu| - \Delta) \end{aligned} \tag{H.13}$$

The other tensor elements are again related by the symmetry transformations of the crystal group,

$$F_A^{yy} = F_A^{xx} \qquad F_A^{xy} = F_A^{yx} = 0 \tag{H.14}$$

Due to time-reversal symmetry, $F_B^{\beta\alpha_1} = 0$, as can be checked by direct evaluation.

The case of the third order conductivity is entirely analogous, but there are, of course, more tensor components and more integrals to assess now,

$$\mathcal{I}_1^{xxyy}(\omega) = \mathcal{I}_1^{xyxy}(\omega) = -\frac{\hbar^2 v_F^2}{2\pi(\hbar\omega)^2}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.15}$$

$$\mathcal{I}_1^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{2\pi(\hbar\omega)^2}\left(1-2\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.16}$$

$$\mathcal{I}_2^{xxyy}(\omega) = \mathcal{I}_2^{xyxy}(\omega) = \mathcal{I}_2^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{4\pi(\hbar\omega)^2}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.17}$$

$$\mathcal{I}_3^{xxyy}(\omega) = \frac{\hbar^2 v_F^2}{4\pi\hbar\omega}\left(3-2\left(\frac{\Delta}{\hbar\omega}\right)^2-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.18}$$

$$\mathcal{I}_3^{xyxy}(\omega) = \mathcal{I}_3^{xyyx}(\omega) = -\frac{\hbar^2 v_F^2}{4\pi\hbar\omega}\left(1-2\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.19}$$

$$\mathcal{I}_4^{xxyy}(\omega) = \mathcal{I}_4^{xyxy}(\omega) = -\frac{2\,\hbar^2 v_F^2}{\pi(\hbar\omega)^3}\left(\frac{\Delta}{\hbar\omega}\right)^4\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.20}$$

$$\mathcal{I}_4^{xyyx}(\omega) = \frac{2\,\hbar^2 v_F^2}{\pi(\hbar\omega)^3}\left(\left(\frac{\Delta}{\omega}\right)^2-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.21}$$

$$\mathcal{I}_5^{xxyy}(\omega) = \frac{\hbar^2 v_F^2}{16\pi(\hbar\omega)^3}\left(1-10\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.22}$$

$$\mathcal{I}_5^{xyxy}(\omega) = \mathcal{I}_5^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{16\pi(\hbar\omega)^3}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.23}$$

$$\mathcal{I}_6^{xxyy}(\omega) = -\frac{\hbar^2 v_F^2}{\pi(\hbar\omega)^3}\left(1-2\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.24}$$

$$\mathcal{I}_6^{xyxy}(\omega) = \mathcal{I}_6^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{\pi(\hbar\omega)^3}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\Theta(\hbar|\omega|-\Delta_{eff}) \tag{H.25}$$

$$\mathcal{I}_j^{xxxx}(\omega) = \mathcal{I}_j^{xxyy}(\omega) + \mathcal{I}_j^{xyxy}(\omega) + \mathcal{I}_j^{xyyx}(\omega) \tag{H.26}$$

and the respective Hilbert transforms,

$$\begin{aligned}\mathcal{H}_1^{xxyy}(\omega) = \mathcal{H}_1^{xyxy}(\omega) &= -\frac{\hbar^2 v_F^2}{2\pi^2(\hbar\omega)^2}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right| \\ &- \frac{\hbar^2 v_F^2}{\pi^2\,\hbar\omega\,\Delta_{eff}}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{\hbar^2 v_F^2\,\Delta^4}{3\,\pi^2\,(\hbar\omega)^3\,\Delta_{eff}^3}+\frac{\hbar^2 v_F^2\,\Delta^4}{5\,\pi^2\,\hbar\omega\,\Delta_{eff}^5}\end{aligned} \tag{H.27}$$

$$\begin{aligned}\mathcal{H}_1^{xyyx}(\omega) &= \frac{\hbar^2 v_F^2}{2\pi^2(\hbar\omega)^2}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right| \\ &+ \frac{\hbar^2 v_F^2}{\pi^2\,\hbar\omega\,\Delta_{eff}}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2-\frac{\hbar^2 v_F^2\,\Delta^2}{3\,\pi^2\,\hbar\omega\,\Delta_{eff}^3}\left(2-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)+\frac{\hbar^2 v_F^2\,\Delta^4}{5\,\pi^2\,\hbar\omega\,\Delta_{eff}^5}\end{aligned} \tag{H.28}$$

$$\mathcal{H}_2^{xxyy}(\omega) = \frac{\hbar^2 v_F^2}{4\pi^2(\hbar\omega)^2}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right| + \frac{\hbar^2 v_F^2\,\Delta^4}{10\,\pi^2\,\hbar\omega\,\Delta_{eff}^5}$$
$$+\frac{\hbar^2 v_F^2}{2\pi^2\,\hbar\omega\,\Delta_{eff}}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{\hbar^2 v_F^2\,\Delta^2}{6\,\pi^2\,\hbar\omega\,\Delta_{eff}^3}\left(6+\left(\frac{\Delta}{\hbar\omega}\right)^2\right) \tag{H.29}$$

$$\mathcal{H}_2^{xyxy}(\omega) = \mathcal{H}_2^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{4\pi^2(\hbar\omega)^2}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$+\frac{\hbar^2 v_F^2}{2\,\pi^2\,\hbar\omega\,\Delta_{eff}}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2-\frac{\hbar^2 v_F^2\,\Delta^2}{6\,\pi^2\,\hbar\omega\,\Delta_{eff}^3}\left(2-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)+\frac{\hbar^2 v_F^2\,\Delta^4}{10\,\pi^2\,\hbar\omega\,\Delta_{eff}^5} \tag{H.30}$$

$$\mathcal{H}_3^{xxyy}(\omega) = \frac{\hbar^2 v_F^2}{4\pi^2\hbar\omega}\left(3-2\left(\frac{\Delta}{\hbar\omega}\right)^2-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$-\frac{\hbar^2 v_F^2\,\Delta^2}{2\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(2+\left(\frac{\Delta}{\hbar\omega}\right)^2\right)-\frac{\hbar^2 v_F^2\,\Delta^4}{6\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3} \tag{H.31}$$

$$\mathcal{H}_3^{xyxy}(\omega) = \mathcal{H}_3^{xyyx}(\omega) = -\frac{\hbar^2 v_F^2}{4\pi^2\hbar\omega}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$+\frac{\hbar^2 v_F^2\,\Delta^2}{2\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(2-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)-\frac{\hbar^2 v_F^2\,\Delta^4}{6\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3} \tag{H.32}$$

$$\mathcal{H}_4^{xxyy}(\omega) = \mathcal{H}_4^{xyxy}(\omega) = -\frac{2\,\hbar^2 v_F^2}{\pi^2\,(\hbar\omega)^3}\left(\frac{\Delta}{\hbar\omega}\right)^4\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$-\frac{4\,\hbar^2 v_F^2\,\Delta^4}{\pi^2\,(\hbar\omega)^6\,\Delta_{eff}}-\frac{20\,\hbar^2 v_F^2\,\Delta^4}{15\,\pi^2\,(\hbar\omega)^4\,\Delta_{eff}^3}-\frac{12\,\hbar^2 v_F^2\,\Delta^4}{15\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5} \tag{H.33}$$

$$\mathcal{H}_4^{xyyx}(\omega) = \frac{2\,\hbar^2 v_F^2}{\pi^2\,(\hbar\omega)^3}\left(\left(\frac{\Delta}{\hbar\omega}\right)^2-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|-\frac{12\,\hbar^2 v_F^2\,\Delta^4}{15\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5}$$
$$+\frac{4\,\hbar^2 v_F^2}{\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(\left(\frac{\Delta}{\hbar\omega}\right)^2-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{20\,\hbar^2 v_F^2\,\Delta^2}{15\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right) \tag{H.34}$$

$$\mathcal{H}_5^{xxyy}(\omega) = \frac{\hbar^2 v_F^2}{16\pi^2(\hbar\omega)^3}\left(1-10\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|+\frac{\hbar^2 v_F^2\,\Delta^4}{40\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5}$$
$$+\frac{\hbar^2 v_F^2}{8\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(1-10\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)-\frac{\hbar^2 v_F^2\,\Delta^2}{24\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3}\left(10-\left(\frac{\Delta}{\hbar\omega}\right)^2\right) \tag{H.35}$$

$$\mathcal{H}_5^{xyxy}(\omega) = \mathcal{H}_5^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{16\pi^2(\hbar\omega)^3}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$+\frac{\hbar^2 v_F^2}{8\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(1+6\left(\frac{\Delta}{\hbar\omega}\right)^2+\left(\frac{\Delta}{\hbar\omega}\right)^4\right)+\frac{\hbar^2 v_F^2\,\Delta^2}{24\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3}\left(6+\left(\frac{\Delta}{\hbar\omega}\right)^2\right)$$
$$+\frac{\hbar^2 v_F^2\,\Delta^4}{40\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5} \tag{H.36}$$

$$\mathcal{H}_6^{xxyy}(\omega) = -\frac{\hbar^2 v_F^2}{\pi^2(\hbar\omega)^3}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right|$$
$$-\frac{2\,\hbar^2 v_F^2}{\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)^2+\frac{10\,\hbar^2 v_F^2\,\Delta^2}{15\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^3}\left(2-\left(\frac{\Delta}{\hbar\omega}\right)^2\right)-\frac{2\,\hbar^2 v_F^2\,\Delta^4}{5\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5} \tag{H.37}$$

$$\mathcal{H}_6^{xyxy}(\omega) = \mathcal{H}_6^{xyyx}(\omega) = \frac{\hbar^2 v_F^2}{\pi^2(\hbar\omega)^3}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right)\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right| + \frac{2\,\hbar^2 v_F^2}{\pi^2\,(\hbar\omega)^2\,\Delta_{eff}}\left(1-\left(\frac{\Delta}{\hbar\omega}\right)^4\right) - \frac{10\,\hbar^2 v_F^2\,\Delta^4}{15\,\pi^2\,(\hbar\omega)^4\,\Delta_{eff}^3} - \frac{2\,\hbar^2 v_F^2\,\Delta^4}{5\,\pi^2\,(\hbar\omega)^2\,\Delta_{eff}^5} \tag{H.38}$$

$$\mathcal{H}_j^{xxxx}(\omega) = \mathcal{H}_j^{xxyy}(\omega) + \mathcal{H}_j^{xyxy}(\omega) + \mathcal{H}_j^{xyyx}(\omega) \tag{H.39}$$

In addition to the listed integrals, an identical set is obtained exchanging the indices x and y. All other tensor elements of the integrals $\mathcal{I}_j^{\beta\alpha_1\alpha_2\alpha_3}$ and $\mathcal{H}_j^{\beta\alpha_1\alpha_2\alpha_3}$ that have not been explicitly mentioned are identically zero, as a consequence of the spatial symmetries of the crystal.

For the passage from the $\mathcal{I}$ to the $\mathcal{H}$ integrals, it is useful to hold in mind the following result on Hilbert transforms[1], valid for $n > 1$,

$$-\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{\Theta(|\hbar\omega'|-\Delta_{eff})}{(\hbar\omega')^n(\omega'-\omega)}d\omega' = \frac{1}{\pi(\hbar\omega)^n}\ln\left|\frac{\hbar\omega-\Delta_{eff}}{\hbar\omega+\Delta_{eff}}\right| + \frac{2}{\pi}\sum_{k=1}^{n/2}\frac{1}{(2k-1)\,\Delta_{eff}^{2k-1}\,(\hbar\omega)^{n+1-2k}} \tag{H.40}$$

Finally, the integrals in the Fermi surface contribution,

$$F_A^{xxxx}/3 = F_A^{xxyy} = F_A^{xyxy} = F_A^{xyyx} = \frac{\hbar^2 v_F^2}{4\pi|\mu|}\left(1+2\left(\frac{\Delta}{2\,\mu}\right)^2-3\left(\frac{\Delta}{2\,\mu}\right)^4\right)\Theta(2|\mu|-\Delta) \tag{H.41}$$

As expected, the exact same result is obtained for $F_A^{yyyy}/3 = F_A^{yyxx} = F_A^{yxyx} = F_A^{yxxy}$, with all other tensor elements being identically zero. As in linear order (and every odd order, in fact), the $F_B^{\beta\alpha_1\alpha_2\alpha_3}$ integrals vanish, due to presence of time-reversal symmetry.

The formulas become simpler when we close the gap $\Delta = 0$. This case can be quickly derived from the above expressions and is in agreement with previously published results. Integral evaluation for monolayer graphene with no band gap is listed in Appendix D of [91].

Replacement of the integrals computed here in Eqs. 5.31-5.34 of Section 5.2.3 reproduces the expression for the third order conductivity derived by Cheng et al. [76] for $\Delta = 0$.

[1] In fact, the Hilbert transform is undefined, diverging for $\pm\infty$. An appropriate result, as in Eq. H.40, can, however, be recovered by introducing a limited bandwidth with an additional Heaviside step function $\Theta(\hbar|\omega| - \epsilon_{\max})$, where the maximum energy is sufficiently removed from the considered frequencies and the effective gap: $\epsilon_{\max} \gg \omega \wedge \epsilon_{\max} \gg \Delta_{eff}$.

List of Figures

List of Tables

Bibliography

[1] J. C. Maxwell, *A treatise on electricity and magnetism* (Oxford: Clarendon Press, 1873).

[2] B. J. Hunt, Physics Today **65** (2012).

[3] P. N. Butcher and D. Cotter, *Elements of nonlinear optics* (Cambridge university press, 1991).

[4] E. Garmire, Optics Express **21** (2013).

[5] J. E. Sipe and A. I. Shkrebtii, Physical Review B **61** (2000).

[6] B. M. Fregoso, R. A. Muniz, and J. E. Sipe, Physical Review Letters **121** (2018).

[7] G. B. Ventura, D. J. Passos, J. M. Viana Parente Lopes, and J. M. B. Lopes dos Santos, Physical Review Letters **126** (2021).

[8] G. B. Ventura, D. J. Passos, J. M. V. Lopes, and J. M. B. Lopes dos Santos, arXiv (2020).

[9] Y.-R. Shen, *The principles of nonlinear optics* (Wiley, 1984).

[10] R. Boyd, *Nonlinear optics* (Academic Press, 2008).

[11] E. Garmire, American Journal of Physics **79** (2011).

[12] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Physical Review Letters **7** (1961).

[13] T. H. Maiman, Physical Review Letters **4** (1960).

[14] T. H. Maiman, Nature (1960).

[15] N. Bloembergen, IEEE Journal on Selected Topics in Quantum Electronics **6** (2000).

[16] W. Kaiser and C. G. B. Garrett, Physical Review Letters **7** (1961).

[17] D. A. Kleinman, Physical Review **125** (1962).

[18] M. Bass, P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, Physical Review Letters **8** (1962).

[19] M. Bass, P. A. Franken, and J. F. Ward, Physical Review **138** (1965).

[20] R. W. Terhune, P. D. Maker, and C. M. Savage, Physical Review Letters **8** (1962).

[21] J. A. Giordmaine, Physical Review Letters **8** (1962).

[22] P. D. Maker, R. W. Terhune, M. Nisenoff, and C. M. Savage, Physical Review Letters **8** (1962).

[23] J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, Physical Review **127** (1962).

[24] D. A. Kleinman, Physical Review **126** (1962).

[25] P. N. Butcher and T. P. McLean, Proceedings of the Physical Society (1963).

[26] R. C. Miller, D. A. Kleinman, and A. Savage, Physical Review Letters **11** (1963).

[27] P. N. Butcher and T. P. Mclean, Proceedings of the Physical Society **83** (1964).

[28] J. M. Manley and H. E. Rowei, Proceedings of the IRE **44** (1956).

[29] H. A. Haus, IRE Transactions on microwave theory and techniques **6** (1958).

[30] D. J. Moss, J. E. Sipe, and H. M. Van Driel, Physical Review B **36** (1987).

[31] E. Garmire, in *Nonlinear Optics* (InTech, 2012).

[32] D. J. Moss, E. Ghahramani, J. E. Sipe, and H. M. Van Driel, Physical Review B **41** (1990).

[33] R. B. Miles and S. E. Harris, IEEE Journal of Quantum Electronics **9** (1973).

[34] C.-L. Pan, C.-Y. She, W. Fairbank, and K. Billman, IEEE Journal of Quantum Electronics **13** (1977).

[35] J. Morrell and A. C. Albrecht, Chemical Physics Letters **64** (1979).

[36] S. J. Lalama and A. F. Garito, Physical Review A **20** (1979).

[37] C. C. Teng and A. F. Garito, Physical Review Letters **50** (1983).

[38] C. C. Teng and A. F. Garito, Physical Review B **28** (1983).

[39] R. K. Chang, J. Ducuing, and N. Bloembergen, Physical Review Letters **15** (1965).

[40] C. Y. Fong and Y. R. Shen, Physical Review B **12** (1975).

[41] D. J. Moss, H. M. Van Driel, and J. E. Sipe, Optics Letters **14** (1989).

[42] E. Ghahramani, D. J. Moss, and J. E. Sipe, Physical Review B **43** (1991).

[43] J. L. P. Hughes and J. E. Sipe, Physical Review B **55** (1996).

[44] J. L. P. Hughes, Y. Wang, and J. E. Sipe, Physical Review B **55** (1997).

[45] D. E. Aspnes, Physical Review B **6** (1972).

[46] K. Rzążewski and R. W. Boyd, Journal of Modern Optics **51** (2004).

[47] J. E. Sipe and E. Ghahramani, Physical Review B **48** (1993).

[48] C. Aversa, J. E. Sipe, M. Sheik-Bahae, and E. W. Van Stryland, Physical Review B **50** (1994).

[49] C. Aversa and J. E. Sipe, Physical Review B **52** (1995).

[50] V. N. Genkin and P. M. Mednis, Soviet Physics JETP **27** (1968).

[51] Z. H. Levine, Physical Review B **49** (1994).

[52] A. Dal Corso and F. Mauri, Physical Review B **50** (1994).

[53] A. Dal Corso and F. Mauri, Physical Review B **53** (1996).

[54] E. I. Blount (Academic Press, 1962).

[55] H. M. Van Driel, J. E. Sipe, A. Haché, and R. Atanasov, Physica Status Solidi (B) Basic Research **204** (1997).

[56] D. Passos, G. Ventura, J. Viana Parente Lopes, J. Lopes Dos Santos, and N. Peres, Physical Review B **97** (2018).

[57] G. Ventura, D. Passos, J. Lopes Dos Santos, J. Viana Parente Lopes, and N. Peres, Physical Review B **96** (2017).

[58] N. W. Ashcroft and N. D. Mermin, *Solid state physics* (Saunders College Publishing, 1976).

[59] M. V. Berry, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **392** (1984).

[60] D. Vanderbilt, *Berry Phases in Electronic Structure Theory* (Cambridge university press, 2018).

[61] R. Kubo, Journal of the Physical Society of Japan **12** (1957).

[62] T. Holder, D. Kaplan, and B. Yan, Physical Review Research **2** (2020).

[63] A. K. Geim and I. V. Grigorieva, Nature **499** (2013).

[64] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, Science **306** (2004).

[65] A. H. Castro Neto, F. Guinea, N. M. Peres, K. S. Novoselov, and A. K. Geim, Reviews of Modern Physics **81** (2009).

[66] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, Nature **438** (2005).

[67] A. A. Balandin, S. Ghosh, W. Bao, I. Calizo, D. Teweldebrhan, F. Miao, and C. N. Lau, Nano Letters **8** (2008).

[68] A. K. Geim, Science **324** (2009).

[69] K. S. Novoselov, V. I. Falko, L. Colombo, P. R. Gellert, M. G. Schwab, and K. Kim, Nature **490** (2012).

[70] R. R. Nair, P. Blake, A. N. Grigorenko, K. S. Novoselov, T. J. Booth, T. Stauber, N. M. Peres, and A. K. Geim, Science **320** (2008).

[71] N. M. Peres, F. Guinea, and A. H. Castro Neto, Physical Review B **73** (2006).

[72] S. A. Mikhailov, Europhysics Letters **79** (2007).

[73] E. Hendry, P. J. Hale, J. Moger, A. K. Savchenko, and S. A. Mikhailov, Physical Review Letters **105** (2010).

[74] S. A. Mikhailov and K. Ziegler, Journal of Physics Condensed Matter **20** (2008).

[75] S. A. Mikhailov, Physical Review B **93** (2016).

[76] J. L. Cheng, N. Vermeulen, and J. E. Sipe, New Journal of Physics **18** (2014).

[77] J. L. Cheng, N. Vermeulen, and J. E. Sipe, Physical Review B **91** (2015).

[78] T. Jiang, D. Huang, J. Cheng, X. Fan, Z. Zhang, Y. Shan, Y. Yi, Y. Dai, L. Shi, K. Liu, C. Zeng, J. Zi, J. E. Sipe, Y. R. Shen, W. T. Liu, and S. Wu, Nature Photonics **12** (2018).

[79] M. I. Katsnelson, Materials Today (2007).

[80] J. Cayssol, Comptes Rendus Physique **14** (2013).

[81] S. Y. Zhou, G. H. Gweon, A. V. Fedorov, P. N. First, W. A. De Heer, D. H. Lee, F. Guinea, A. H. Castro Neto, and A. Lanzara, Nature Materials **6** (2007).

[82] C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, and J. Hone, Nature Nanotechnology **5** (2010).

[83] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, 1995).

[84] G. Ventura, D. Passos, J. Viana Parente Lopes, and J. Lopes Dos Santos, Journal of Physics Condensed Matter **32** (2020).

[85] D. E. Parker, T. Morimoto, J. Orenstein, and J. E. Moore, Physical Review B **99** (2019).

[86] S. M. João and J. M. Viana Parente Lopes, Journal of Physics Condensed Matter **32** (2020).

[87] S. M. João and J. M. Viana Parente Lopes, EPJ Web of Conferences **233** (2020).

[88] A. Jorio and M. Dresselhaus, *Group theory: application to the physics of condensed matter* (Springer Science & Business Media, 2007).

[89] M. I. Katsnelson, *Graphene: carbon in two dimensions* (Cambridge university press, 2012).

[90] S. Y. Hong, J. I. Dadap, N. Petrone, P. C. Yeh, J. Hone, and R. M. Osgood, Physical Review X **3** (2013).

[91] D. J. Passos, G. B. Ventura, J. M. B. Lopes dos Santos, and J. M. Viana Parente Lopes, Journal of Physics: Condensed Matter **33** (2021).

[92] L. Allen and J. H. Eberly, *Optical resonance and two-level atoms* (Courier corporation, 1987).

[93] R. Resta, *Geometry and topology in electronic structure theory* (2015).

[94] J. L. Cheng, J. E. Sipe, S. W. Wu, and C. Guo, APL Photonics **4** (2019).

[95] D. C. Hutchings, M. Sheik-Bahae, D. J. Hagan, and E. W. Van Stryland, Optical and Quantum Electronics **24** (1992).

[96] G. B. Ventura, *Nonlinear optical response of two-dimensional crystals*, Ph.D. thesis, University of Porto (2021).

[97] D. Kaplan, T. Holder, and B. Yan, arXiv (2020).

[98] E. V. Castro, K. S. Novoselov, S. V. Morozov, N. M. R. Peres, J. Nilsson, A. K. Geim, and A. H. C. Neto, Physical Review Letters **99** (2007).

[99] T. Morimoto and N. Nagaosa, Science Advances **2** (2016).

[100] L. Wu, S. Patankar, T. Morimoto, N. L. Nair, E. Thewalt, A. Little, J. G. Analytis, J. E. Moore, and J. Orenstein, Nature Physics **13** (2017).

[101] G. B. Osterhoudt, L. K. Diebel, M. J. Gray, X. Yang, J. Stanco, X. Huang, B. Shen, N. Ni, P. J. Moll, Y. Ran, and K. S. Burch, Nature Materials **18** (2019).

[102] B. M. Fregoso, Physical Review B **100** (2019).

[103] J. L. Cheng, J. E. Sipe, and C. Guo, Physical Review B **10** (2019).

[104] L. Yang, J. Deslippe, C. H. Park, M. L. Cohen, and S. G. Louie, Physical Review Letters **103** (2009).