Politecnico di Torino

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation
Doctoral Program in Electrical, Electronics and Communications Engineering (35th cycle)

Full-band NEGF modeling of optoelectronic devices

Presented by:

Jesús Alberto González Montoya

******

Supervisors:

Prof. Francesco Bertazzi, Supervisor
Prof. Michele Goano, Co-Supervisor
Dr. Alberto Tibaldi, Co-Supervisor

Doctoral Examination Committee:

Prof. David Esseni, University of Udine (Udine, Italy)
Prof. Mathieu Luisier, ETH (Zurich, Switzerland)
Prof. Tillmann Kubis, Purdue University (Indiana, USA)
Prof. Simona Donati Guerrieri, Politecnico di Torino (Torino, Italy)
Prof. Fabrizio Dolcini, Politecnico di Torino (Torino, Italy)

Politecnico di Torino
June 2023

Acknowledgments

First of all, I want to start by thanking God for giving me the opportunity to conclude this important stage of my life. Second, I would like to dedicate this work to all my family, my friends and specially, to my sisters and my mother, which in spite of the distance have provided me with invaluable support, belief and love, and have been one of my main motivations during my academic career. Similarly, I want to thank my beloved wife Camila for being at my side all these three years, for her endless patience and love, and for being one of my main emotional supports in the good and bad times.

Next, I would like to express my most sincere gratitude to my advisor Prof. Francesco Bertazzi for offering me the opportunity to work in such an interesting field, for his continuous support and patience, and for the great liberty he gave me during its realization. He was always available to discuss my work and his guidance helped me during all the time of research and the writing of this thesis.

I would also like to offer my deepest gratitude to Prof. Michele Goano for his continuous support in all the problems I encountered out of the academic framework, and for his friendly and valuable advices. I am particularly grateful to him and Prof. Bertazzi for providing me financial support during the last few months of the thesis, and for offering me the research position I currently have. I am also indebted to Dr. Alberto Tibaldi for his continuous support and availability. He provided me with the most useful comments, observations and suggestions on my research and during the writing of the thesis.

Finally, I wish to extend my gratitude to the Politecnico di Torino for the financial and formative support, to the Department of Electronics and Telecommunications (DET) and to all my colleges in the Microwave and Optoelectronics Group (MOG). Special thanks also to Prof. Matteo Meneghini and his co-workers from the Università di Padova for the helpful discussions and for the experimental data they provided me.

Abstract

Theoretical investigation of devices at the nanoscale should take into account several quantum effects that dominate transport and optical properties at these dimensions. Research on nanostructures, which has been mostly limited to simplified semiclassical treatments with quantum corrections and *ad-hoc* models for the various effects, still seems not satisfactory in many devices where quantum properties are actively exploited. Consequently, quantum-kinetic models continue to acquire greater relevance in the field of numerical simulation, despite their enormous computational cost, which puts stringent restrictions to the geometrical dimensions of the structures under investigation and the band structure model.

In this thesis we aim to provide a deep insight into the relevant aspects of the nonequilibrium Green's function (NEGF) formalism and its application to the numerical simulation of optoelectronic devices, specifically oriented to LEDs and photodetectors. Our implementation is based on a multiband $\mathbf{k} \cdot \mathbf{p}$ representation of the electronic dispersion, which is essential for a realistic description of intra or interband tunneling processes and optical transitions. Carrier dynamics are obtained by solving the Dyson's equation of motion in steady-state conditions, and particle interactions are included through the corresponding scattering self-energies computed in the self-consistent Born approximation (SCBA). Microscopic quantities such as the carrier densities, the local density of states and the scattering rates, are obtained in NEGF with energy-resolution, allowing a profound understanding of the underlying physical phenomena.

The present model, based on a finite-element space discretization in the longitudinal direction, is applied to the analysis of two nanostructured devices of particular relevance: GaN-based light emitting diodes (LED) and superlattice infrared photodetectors. In the first case, we investigate carrier transport in the subthreshold forward-bias regime where tunneling processes are relevant. The presence of defects has been often correlated with the high ideality factors experimentally observed in the subthreshold I-V characteristics of III-nitride devices. A quantum kinetic theory of trap-assisted tunneling is developed within the NEGF framework, and we demonstrate that the quantum nonradiative recombination rates can be reproduced by the conventional Shockley-Read-Hall theory, provided that the classical charge is replaced with the correct quantum charge. Trap-assisted tunneling can be

described with drift-diffusion solvers complemented with appropriate quantum corrections for the calculation of the local density of states. The proposed model is shown to predict the subthreshold I-V characteristics and ideality factors from experimental data taken on a single quantum-well LED.

As a second case study, we analyze the transport properties of superlattice infrared absorbers. Inspection of spatially and energetically resolved single particle properties offers insight into the complex nature of carrier transport in type-II superlattice detectors. The use of a genuine quantum-kinetic model is crucial here for an accurate description of the main ingredients determining the different transport regimes (*i.e.*, miniband transport, sequential tunneling, hopping). The computational challenges posed by the model makes it necessary to keep a rather simple approach, limiting the geometrical sizes of the simulated structures. Moreover, a large number of iterations is needed to achieve self-consistency between Green's functions and self-energies in the presence of strongly localized states not directly accessible from the contacts. We attempt to demonstrate in this work an accurate, yet computationally feasible NEGF model of superlattice detectors by formulating the kinetic equations in terms of problem-matched maximally localized basis functions, numerically generated from few modes representing the main conductive channels of the nanostructure. A mobility study is also performed, providing a semiclassical interpretation of the obtained numerical results.

Content

Acknowledgments iii

Abstract iv

1 Introduction 3

1.1 Motivation . . . 3

1.2 Outline . . . 5

2 Nonequilibrium Green's Function device modeling 8

2.1 Contour-ordered nonequilibrium Green's function . . . 9

2.1.1 Real-time analytic continuations of the Green's function . . . 12

2.2 Green's functions in energy space . . . 14

2.3 Physical observables . . . 16

2.3.1 Carrier densities . . . 16

2.3.2 Local density of states . . . 17

2.3.3 Current density . . . 17

2.4 Spatial discretization . . . 20

2.5 Choice of basis: the $\mathbf{k} \cdot \mathbf{p}$ model . . . 23

2.5.1 The $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian for layered devices . . . 27

2.6 Boundary conditions and boundary self-energy . . . 33

2.6.1 Retarded boundary self-energy of the contacts . . . 35

2.7 Particle interactions . . . 38

2.7.1 Electron-phonon self-energy . . . 42

2.7.2 Electron-photon self-energy . . . 46

2.7.3 Carrier-carrier self-energy . . . 48

2.7.4 Büttiker probes . . . 54

2.8 Poisson equation . . . 55

2.8.1 Ionized dopant concentrations . . . 57

2.8.2 Polarization charges . . . 57

2.9 Numerical considerations and choice of simulation parameters . . . 58

3 Modeling of trap-assisted tunneling in GaN based LEDs **62**
3.1 NEGF treatment of localized deep defects 62
3.1.1 Detailed balance principle of the defect recombination rate 66
3.1.2 Derivation of the multiphonon matrix elements 69
3.2 Shockley-Read-Hall self-energy 72
3.3 Enhancement of the defect recombination by tunneling 75
3.4 Simulation of an InGaN/GaN LED 83
3.4.1 Numerical considerations 84
3.4.2 Defect recombination in the subthreshold regime 85
3.4.3 NEGF connection with the semiclassical SRH theory 89
3.4.4 Fitting of the I-V characteristic 91

4 NEGF modeling of type-II superlattice detectors **94**
4.1 Brief introduction to type-II superlattice infrared detectors 95
4.2 Mode-space analysis 98
4.3 Accuracy assessment of the mode-space approach on an ICIP device 100
4.4 Connection to semiclassical approaches: mobility study in a LWIR type-II SL 107
4.5 Transport across a barrier in a superlattice 114
4.5.1 Numerical considerations 114
4.5.2 Simulation results 119

5 Conclusions and Outlook **124**

Appendix A The 8-band $\mathbf{k}\cdot\mathbf{p}$ Luttinger-Kohn Hamiltonian **129**
A.1 Zinc-blende model 129
A.2 Wurtzite model 131

Appendix B Carrier interaction with photons and phonons **135**
B.1 Interaction Hamiltonian and perturbation expansion 135
B.2 Derivation of the self-energy 138
B.2.1 Renormalization of the boson Green's function 140

Bibliography **144**

List of publications **155**

Chapter 1

Introduction

1.1 Motivation

Semiconductor electronic devices for light generation and detection applications have gained a tremendous relevance over the last decades. With a wide range of applications in nearly any field, from general purpose lighting, energy harvesting, to optical sensors for food, surveillance or environmental control, it is undeniable the role that solid-state optoelectronics play in our current society. Moreover, the increment in the global demand of electrical energy and the constant push towards renewable energy sources has driven researchers to focus on the development of technologies with increased efficiency and reduced operational current. In this context, numerical modeling plays a very important part in the optimization of device performance and the understanding of the underlying physics dominating the device operation.

In general, theoretical investigation in the field of optoelectronics has been limited to semi-classical transport models derived from the *Boltzmann transport equation* (BTE), and in particular to the *drift-diffusion* (DD) model, which is obtained from the zeroth-order moment of the BTE. Carrier interaction with electromagnetic waves is included in the form of generation-recombination rates of electron-hole pairs [1], and the dynamic evolution of the light field is modeled through an optical solver coupled to the electron system in the device by absorption/gain coefficients. Nevertheless, nanostructured material systems, which compose the active region of any modern optoelectronic device, are highly affected by the quantum nature of carriers, which could be engineered as, *e.g.*, in the case of carrier confinement in light-emitting diodes (LEDs), or tunneling in resonant tunneling diodes (RTDs) and tunneling-filed-effect transistors (TFETs). Thus, adequate quantum corrections should be added to the above picture in order take into account effects like the coupling of the extended bulk states (responsible of the current flow through the device) to the confined states, and tunneling-assisted recombination processes.

At this point, the question remains about the appropriateness of the models used for the quantum corrections, and even more, if the quantum corrections themselves within a semiclassical framework are sufficient to describe all the complex dynamics that interplay in the physics of the nanostructure. On the one hand, semiclassical models like DD implicitly assume an effective mass approximation of the electronic dispersion with local interactions in space, so hot carriers and non-local effect are out of the picture. On the other hand, there are open debates about the loss mechanisms taking place, mainly because of the slight degree of ambiguity left in the parameters of the ABC model, which leave unclear to what extent each recombination process intervene at some given conditions. Consider for example the case of GaN-based LEDs and the *droop* controversy, *i.e.*, the decline of the internal quantum efficiency experimentally observed at high current densities [2, 3, 4]. Recent experimental studies on droop suggest that several physical mechanisms may contribute to this effect, such as alloy- and trap-assisted Auger [5, 6, 7], or density-activated defect assisted recombination [8]. These are complex processes which can be hardly described well enough in terms of simple expressions, or without requiring to introduce an additional set of parameters whose values are difficult to estimate experimentally. Even more, reproducing macroscopic experimental data within a semiclassical framework with an appropriate set of fitting parameters does not imply that the model is realistic at the microscales. The fundamental problems arising from the inclusion of nonclassical correction within a semiclassical description emphasizes the necessity of a unified framework in which optical and transport aspects are treated on equal footing and quantum effects can be directly included in the recombination processes.

Overcoming the semiclassical description entails moving forward to genuine quantum kinetic approaches, able to provide a reliable approximation to the numerical simulation of optoelectronic devices, at least in what concerns their active region, and a thorough understanding of the physical mechanisms behind the microscopic phenomena with a high degree of fidelity. Quantum transport techniques come into play thanks to the development of theories like nonequilibrium Green's function (NEGF) formalism [9] and density matrix (DM) theory [10], suitable for the modeling of nanostructured devices in steady state and transient regimes, respectively. Leonid Keldysh in 1965 [11] firstly presented his theory of NEGF and opened the gate for this formalism to become a real computational tool. Recent technological breakthroughs permitted to produce full-operative complex optoelectronic devices like quantum cascade lasers (QCL) or vertical-cavity surface-emitting lasers (VCSEL), which brought the need of realistic simulation tools. This, together with the recent advent of large computer clusters with high computational capabilities, has given the NEGF approach such great popularity for the numerical simulation of optoelectronic structures, mainly due to its relative simplicity and inherent capability to integrate the electronic transport and the optical domain into a single theory.

In the present work, our main interest will be the study of carrier transport in semiconductor LEDs and infrared photodetectors, using the NEGF approach. In the first case, we will use this theory for the investigation of defect assisted recombination in GaN-based LEDs, which has been pointed out as one of the main causes leading to the droop phenomenon. At low current densities, before the radiative recombination sets in, the role of defects is more evident, as traps may assist the tunneling of carriers, opening additional current leakage channels that would not be accessible by coherent band-to-band tunneling (BTBT). Indeed, the high ideality factors experimentally observed in the subthreshold I-V characteristics of GaN-based LEDs is considered the signature of trap-assisted tunneling (TAT) [12, 13, 14, 15]. Therefore, TAT is a sensitive indicator of the presence of defects, device growth quality, and degradation due to accelerated stress [16]. Moving to the second case, we will be interested in studying the transport properties of type-II superlattice photodetectors, particularly focused in the long-wavelength (LWIR) spectral region. These devices are characterized for presenting an extreme diversity of the possible carrier-transport mechanisms, which range from miniband transport to Wannier-Stark hopping, depending on geometrical parameters, temperature, and the presence of built-in and/or applied fields [17]. Apart from presenting a NEGF study of type-II superlattices, the importance of this work relies on the relevance of this structures as promising alternatives to state-of-the-art infrared detector technology based on mercury cadmium telluride ($Hg_{1-x}Cd_xTe$), with the additional potential advantage of the competitive costs afforded by the standard III-V growth technology.

1.2 Outline

This thesis is divided in four chapters, which are organized as follows: in the second chapter, we deal mainly with presenting the relevant theory concerning the NEGF approach and its numerical implementation. On the one hand, we first present the theoretical aspects of NEGF, in the Keldysh formalism, and we derive expressions for the computation of microscopic quantities such as carrier and current densities, and scattering rates, that will be used in later chapters. Spatial discretization is performed using the finite-element method, with a multiband description of the electronic dispersion based on the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian. Electron interaction with phonons, photons and carriers is described in terms of the corresponding scattering self-energies.

In the third chapter, we are concerned with the study of trap-assisted tunneling in GaN-based LEDs. We start by presenting a theoretical derivation of the self-energy accounting for the interaction of electrons with point defects and we establish a connection with the semiclassical Shockley-Read-Hall theory. After that, the presented theory is used to study defect-mediated recombination in an InGaN single quantum-well LED, where we describe the role of trap-assisted tunneling in the low forward bias regime.

Finally, we move to the study of the main transport mechanisms in LWIR superlattice absorbers. Several structures are investigated, and a connection with semiclassical approaches is done by performing a mobility study on a superlattice. Additionally, we present a mode-space approach in which the NEGF equations are projected into a lower-order basis set of maximally localized basis functions. Apart from demonstrating the accuracy of the proposed method, this approach will permit us to model relatively long structures, with lengths over 100 nm.

Chapter 2

Nonequilibrium Green's Function device modeling

In modern optoelectronic devices, quantum effects are largely exploited in order to achieve desirable properties. Among the most common examples can be found quantum well LEDs using carrier confinement to obtain specific emission or absorption frequencies, infrared detectors based on superlattice structures to achieve tunable cut-off wavelengths, or tunneling junctions used in vertical-cavity surface-emitting lasers (VCSELs) to enhance hole injection in the active region while reducing optical absorption. The simulation of optoelectronic devices featuring nanostructures should then take into account all the relevant quantum processes that occur at the nanoscale and dominate the behavior of the device.

Semiclassical models derived from the Boltzmann transport equation (BTE), among which we can find the *drift-diffusion* (DD) model (zeroth-order moment of the BTE), despite being accurate for a general approximation, still lack the capability to reproduce important quantum phenomena like carrier confinement, tunneling, or carrier interaction with photons. Such effects are usually introduced through external models coupled to the transport equations and involving additional parameters that end up serving in many cases to fit experimental data. In order to improve this deficiencies, rigorous quantum transport approaches based on the exact solution of the Schrödinger equation are essential, providing also the capability to extract important quantities as statistical ensemble averaged properties of the device.

In this chapter, we will review the basic theory concerning the Nonequilibrium Green's Function (NEGF) formalism for device transport simulation including various particle interactions, and we will derive expressions to obtain physical observable quantities such as carrier and current densities, and recombination rates. We will restrict our analysis to the study of layered semiconductors, *i.e.*, devices in which the crystal symmetry is broken along a growth direction, hence allowing to choose a discretization basis along the growth axis (1D simulation) while assuming a homogeneous crystalline structure in the transverse direction.

2.1 Contour-ordered nonequilibrium Green's function

This first section will be devoted to define the nonequilibrium Green's function (abbreviated as GF for simplicity), with times defined over a contour in the complex plane, and derive its dynamical equation, the so called *Dyson's equation*. Let us start by considering a system composed of many particles which undergo complex interactions with themselves and with the surrounding medium by means of different scattering mechanisms involving photons, phonons, impurities, etc. The system is assumed to be composed of a finite region called the *device*, where the main phenomena occur, which is connected at its boundaries to the *contacts*, that are treated as open boundaries, *i.e.*, they are particle reservoirs at thermal equilibrium characterized by a temperature T and a chemical potential μ. Such a system is described by the Hamiltonian

$$\hat{\mathcal{H}}(t) = \hat{H} + \hat{H}^{ext}(t) = \hat{H}_0 + \hat{H}^i + \hat{H}^{ext}(t), \tag{2-1}$$

where the time-independent Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}^i$ describes the isolated system at equilibrium, with $\hat{H}_0$ the exactly solvable part containing the effect of lattice and electrostatic potentials, and $\hat{H}^i$ the part containing all the different particle interactions with photons, phonons, impurities, etc. On the other hand, $\hat{H}^{ext}(t)$ is a time-dependent perturbation that drives the system out of equilibrium, applied at times $t > t_0$, and that could be, *e.g.*, an electric field, a light excitation pulse, or an imbalance of the chemical potentials of the reservoirs coupled to the *device*. The properties of this system can be described by means of the single-particle nonequilibrium GF, which is defined as the nonequilibrium ensemble average of contour-ordered field operators [18]

$$\begin{aligned} G(\mathbf{r}, t; \mathbf{r}', t') &= -\frac{\mathrm{i}}{\hbar} \langle \hat{T}_C \{ \hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t) \hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r}', t') \} \rangle \\ &= -\frac{\mathrm{i}}{\hbar} \left(\theta_C(t, t') \langle \hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t) \hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r}', t') \rangle \mp \theta_C(t', t) \langle \hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r}', t') \hat{\Psi}_{\mathcal{H}}(r, t) \rangle \right), \end{aligned} \tag{2-2}$$

where the function $\theta_C(t, t')$ is the Heaviside step function defined on the contour C,

$$\theta_C(t, t') = \begin{cases} 1 & \text{if } t \text{ later than } t' \text{ on } C \\ 0 & \text{otherwise,} \end{cases} \tag{2-3}$$

and $\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t)$ ($\hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r}', t')$) is the field operator that annihilates (creates) a particle at position $\mathbf{r}$ ($\mathbf{r}'$) and at time t (t'), expressed in the Heisenberg picture. The minus (plus) sign stands for fermions (bosons), and the average is taken over the *grand-canonical ensemble*[1], based on our initial assumptions on the system. Since we are just interested in the steady-state

[1]Here we assume that the reader is familiarized with some concepts of many-body quantum theory and quantum statistical mechanics. For further reading, a detailed description of the NEGF theory can be found in [19, 20, 21, 22].

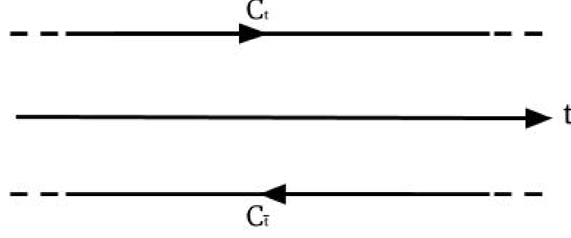


Figure 2-1: *The Schwinger-Keldysh contour, composed of the chronological $C_t = (-\infty, \infty)$ and antichronological $C_{\bar{t}} = (\infty, -\infty)$ branches. Small shift from real axis is just illustrative to distinguish the integration direction.*

condition of the system, time integrals will be taken over the Schwinger-Keldysh contour $C = C_t \cup C_{\bar{t}}$ [23], shown in Fig. **2-1**, in which initial correlations are neglected. A more general treatment of the GF involving transient phenomena can be found in [19] for the interested reader.

The main goal now is to derive the equations of motion of $G(\mathbf{r}, t; \mathbf{r}', t')$ with respect to the times t and t', which we can obtain using the dynamical equations of the field operators in the Heisenberg picture. We will start by considering a Hamiltonian $\hat{\mathcal{H}}$ that includes carrier-carrier interactions $\hat{H}^i = V$ and a single-particle potential $\hat{H}^{ext} = U$, which in second quantization reads

$$\hat{\mathcal{H}}(t) = \int d\mathbf{r}\, \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}, t)[\hat{H}_0(\mathbf{r}) + U(\mathbf{r}, t)]\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t) + \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}'\, \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}, t)\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t)V(\mathbf{r} - \mathbf{r}')\hat{\Psi}_{\mathcal{H}}(\mathbf{r}', t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t). \tag{2-4}$$

This representation of the Hamiltonian is called *normal-ordered*, *i.e.*, all annihilation operators appear to the right of any creation operator. Taking into account the anti-commutation property of the fermionic field operator

$$\left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t), \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t)\right]_+ = \delta(\mathbf{r} - \mathbf{r}'), \tag{2-5}$$

the time derivative of the contour-ordered pair of field operators defined in (2-2) is

$$\begin{aligned} \frac{\partial}{\partial t}\langle \hat{T}_C\{\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t)\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t')\}\rangle &= \left\langle \hat{T}_C\left\{\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t)\right]\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t')\right\}\right\rangle + \delta_C(t, t')\left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t)\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t')\right]_+ \\ &= \left\langle \hat{T}_C\left\{\left[\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(\mathbf{r}, t)\right]\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}', t')\right\}\right\rangle + \delta_C(t, t')\delta(\mathbf{r} - \mathbf{r}'). \end{aligned} \tag{2-6}$$

where $\delta_C(t, t')$ is the derivative with respect to t of the Heaviside function $\theta_C(t, t')$. The time derivative of the annihilation operator can be derived from the corresponding Heisenberg

equation of motion

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) &= \left[\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t),\hat{\mathcal{H}}(t)\right]_{-} \\
&= [\hat{H}_0(\mathbf{r})+U(\mathbf{r},t)]\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t)+\int d\mathbf{r}'V(\mathbf{r}-\mathbf{r}')\hat{\Psi}^{\dagger}_{\mathcal{H}}(\mathbf{r}',t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r}',t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \quad (2\text{-}7)
\end{aligned}
$$

In the following we will use a common short-hand notation in the NEGF theory by writing the Green's function $G(\mathbf{r}_1,t_1;\mathbf{r}_{1'},t_{1'}) = G(11')$, and the contour integral $\int_C d1 = \int_C dt_1 \int d\mathbf{r}_1$. Replacing (2-7) in (2-6) gives the equation of motion of the GF relative to the first time argument t_1 (a similar procedure can be used to derivative Eq. (2-8b))

$$
\left[i\hbar\frac{\partial}{\partial t_1}-\hat{H}_0(\mathbf{r}_1)-U(1)\right]G(11') = \delta(11') - i\hbar\int_C d2V(1-2)G^{(2)}(121'2^+) \quad (2\text{-}8a)
$$

$$
\left[-i\hbar\frac{\partial}{\partial t_{1'}}-\hat{H}_0(\mathbf{r}_{1'})-U(1')\right]G(11') = \delta(11') - i\hbar\int_C d2V(1'-2)G^{(2)}(12^-1'2), \quad (2\text{-}8b)
$$

where we have introduced the quantities

$$
V(1-2) = V(\mathbf{r}_1-\mathbf{r}_2)\delta(t_1-t_2) \quad \text{and} \quad \delta(12) = \delta_C(t_1,t_2)\delta(\mathbf{r}_1-\mathbf{r}_2),
$$

and the two-particle GF, which is defined as follows

$$
G^{(2)}(121'2') = \left(-\frac{i}{\hbar}\right)^2 \left\langle \hat{T}_C\{\hat{\Psi}_{\mathcal{H}}(1)\hat{\Psi}_{\mathcal{H}}(2)\hat{\Psi}^{\dagger}_{\mathcal{H}}(2')\hat{\Psi}^{\dagger}_{\mathcal{H}}(1')\}\right\rangle. \quad (2\text{-}9)
$$

The notation 2^+ in Eq. (2-8a) stands for a time argument in $\hat{\Psi}^{\dagger}_{\mathcal{H}}(2^+)$ infinitesimally larger than t_2, such that the time ordering operator $\hat{T}_C$ in $G^{(2)}$ can reproduce the correct order of the field operators from Eq. (2-7). In the same way, 2^- is intended for a time infinitesimally smaller than t_2. From Eq. (2-8a), it can be seen that the singles-particle GF depends on the two-particle GF $G^{(2)}$, and similar equations can be found for $G^{(2)}$ depending on $G^{(3)}$, $G^{(3)}$ on $G^{(4)}$, and so on, presenting in this way an infinite hierarchy[2] of coupled equations involving GF of increasing order. The difficulties arising from the computation of $G^{(2)}$ are solved by approximating the two-particle GF to different degrees: Hartree-Fock, collision, second Born approximation, etc [19]. In particular, we will follow a different path by defining a quantity known as the self-energy $\Sigma(11')$ encoding interaction effects of all particles on the single-particle dynamics

$$
\left[i\hbar\frac{\partial}{\partial t_1}-\hat{H}_0(\mathbf{r}_1)-U(1)\right]G(11') = \delta(11') + \int_C d3\Sigma(13)G(31') \quad (2\text{-}10a)
$$

$$
\left[-i\hbar\frac{\partial}{\partial t_{1'}}-\hat{H}_0(\mathbf{r}_{1'})-U(1')\right]G(11') = \delta(11') + \int_C d3G(13)\Sigma(31'). \quad (2\text{-}10b)
$$

[2]The so called Martin-Schwinger hierarchy (MSH) of differential equations that couples the n-particle Green's functions of all orders. The initial conditions of this set of equations are given by the Kubo-Martin-Schwinger (KMS) relations and the boundary conditions depend on the physical boundary conditions of the many-body system. See Ref. [21] Sec. 5.1.

Eqs. (2-8a) and (2-10a) (or (2-8b) and (2-10b)) can be demonstrated to be equivalent, and the self-energy may be obtained to different levels of approximations by a perturbative expansion of the GF in (2-2), where the use of a Feynman diagrammatic description of the resulting equations leads to a simple and natural way to arrive to the formal derivation of the self-energy term. Integrating Eqs. (2-10a) and (2-10b), results in the *Dyson's equations*

$$G(11') = G_0(11') + \int_C d2 \int_C d3\, G_0(12)\Sigma(23)G(31') \tag{2-11a}$$

$$G(11') = G_0(11') + \int_C d2 \int_C d3\, G(12)\Sigma(23)G_0(31'), \tag{2-11b}$$

with the *noninteracting Green's function* given by

$$G_0(11') = \left[\mathrm{i}\hbar \frac{\partial}{\partial t_1} - \hat{H}_0(\mathbf{r}_1) - U(1) \right]^{-1} \delta(11'). \tag{2-12}$$

2.1.1 Real-time analytic continuations of the Green's function

The difficulties arising from the complex time-contour integration in the Dyson's equations (2-11a), make it necessary to introduce GFs with real-time arguments as piecewise analytic continuations of $G(11')$

$$G(11') = \begin{cases} G^{t}(11') & t_1, t_{1'} \in C_t \\ G^{\bar{t}}(11') & t_1, t_{1'} \in C_{\bar{t}} \\ G^{<}(11') & t_1 \in C_t, t_{1'} \in C_{\bar{t}} \\ G^{>}(11') & t_1 \in C_{\bar{t}}, t_{1'} \in C_t, \end{cases} \tag{2-13}$$

which are named the *chronological*, *antichronological*, *lesser*, and *greater* Green's functions, respectively, and are defined by

$$G^{t}(11') = -\frac{\mathrm{i}}{\hbar} \langle \hat{T}_t \{ \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^{\dagger}_{\mathcal{H}}(1') \} \rangle \tag{2-14a}$$

$$G^{\bar{t}}(11') = -\frac{\mathrm{i}}{\hbar} \langle \hat{T}_{\bar{t}} \{ \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^{\dagger}_{\mathcal{H}}(1') \} \rangle \tag{2-14b}$$

$$G^{<}(11') = +\frac{\mathrm{i}}{\hbar} \langle \hat{\Psi}^{\dagger}_{\mathcal{H}}(1') \hat{\Psi}_{\mathcal{H}}(1) \rangle \tag{2-14c}$$

$$G^{>}(11') = -\frac{\mathrm{i}}{\hbar} \langle \hat{\Psi}_{\mathcal{H}}(1) \hat{\Psi}^{\dagger}_{\mathcal{H}}(1') \rangle. \tag{2-14d}$$

The chronological and antichronological GFs are usually replaced by the *retarded* and *advanced* Green's functions

$$G^{R}(11') = -\frac{\mathrm{i}}{\hbar}\theta(t_1 - t_{1'}) \langle [\hat{\Psi}_{\mathcal{H}}(1), \hat{\Psi}^{\dagger}_{\mathcal{H}}(1')]_+ \rangle = \theta(t_1 - t_{1'})[G^{>}(11') - G^{<}(11')] \tag{2-15}$$

$$G^{A}(11') = \frac{\mathrm{i}}{\hbar}\theta(t_{1'} - t_1) \langle [\hat{\Psi}_{\mathcal{H}}(1), \hat{\Psi}^{\dagger}_{\mathcal{H}}(1')]_+ \rangle = -\theta(t_{1'} - t_1)[G^{>}(11') - G^{<}(11')]. \tag{2-16}$$

Subtracting the last two equations, we obtain the relation

$$G^R(11') - G^A(11') = G^>(11') - G^<(11'). \tag{2-17}$$

Moreover, due to symmetry properties of G on the contour [24], these four functions satisfy the following relations

$$[G^{\lessgtr}(11')]^\dagger = -G^{\lessgtr}(11') \tag{2-18a}$$
$$[G^R(11')]^\dagger = G^A(11'), \tag{2-18b}$$

which shows that only two of them are linearly independent. The equations of motion for the real-time GFs can be derived by applying the Langreth rules (see [20], Sec. 4.3) to the contour integrals in (2-11a). In particular, a product of the form

$$D(t,t') = \int_C dt_1 \int_C dt_2\, A(t,t_1)B(t_1,t_2)C(t_2,t'),$$

can be decomposed into the real-time integrals

$$D^{R(A)}(t,t') = \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, A^{R(A)}(t,t_1)B^{R(A)}(t_1,t_2)C^{R(A)}(t_2,t') \tag{2-19}$$

$$\begin{aligned} D^{\lessgtr}(t,t') = \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, [&A^R(t,t_1)B^R(t_1,t_2)C^{\lessgtr}(t_2,t') \\ &+ A^R(t,t_1)B^{\lessgtr}(t_1,t_2)C^A(t_2,t') + A^{\lessgtr}(t,t_1)B^A(t_1,t_2)C^A(t_2,t')]. \end{aligned} \tag{2-20}$$

Applying (2-19) to (2-11a) leads to the *Dyson's equations* for the retarded and advanced GFs

$$G^{R(A)}(11') = G_0^{R(A)}(11') + \int d2 \int d3\, G_0^{R(A)}(12)\Sigma^{R(A)}(23)G^{R(A)}(31'), \tag{2-21}$$

with $\int d1 = \int d\mathbf{r}_1 \int_{-\infty}^{\infty} dt_1$. On the other hand, applying (2-20) to (2-11a) for the lesser and greater components gives

$$\begin{aligned} G^{\lessgtr}(11') = G_0^{\lessgtr}(11') + \int d2 \int d3\, \big[&G_0^R(12)\Sigma^R(23)G^{\lessgtr}(31') + G_0^R(12)\Sigma^{\lessgtr}(23)G^A(31') \\ &+ G_0^{\lessgtr}(12)\Sigma^A(23)G^A(31')\big]. \end{aligned} \tag{2-22}$$

Rearranging some terms on the previous equation, it is possible to remove the dependency of (2-22) on $G^{\lessgtr}$. In the following, we adopt momentarily the simplified notation $AB = \int d2\, A(12)B(21')$ where each product imply integration over the internal variables. Eq. (2-22) is then written

$$G^{\lessgtr} = G_0^{\lessgtr} + G_0^R\Sigma^R G^{\lessgtr} + G_0^R\Sigma^{\lessgtr}G^A + G_0^{\lessgtr}\Sigma^A G^A. \tag{2-23}$$

Multiplying each side by $(1+G^R\Sigma^R)$ on the left, and using Eq. (2-21), gives the final result[3]

$$\begin{aligned} G^{\lessgtr} &= (1+G^R\Sigma^R)G_0^{\lessgtr}(1+\Sigma^A G^A) + (G_0^R + G^R\Sigma^R G_0^R)\Sigma^{\lessgtr}G^A \\ &= (1+G^R\Sigma^R)G_0^{\lessgtr}(1+\Sigma^A G^A) + G^R\Sigma^{\lessgtr}G^A. \end{aligned} \tag{2-24}$$

From Eqs. (2-10a)-(2-10b), $G^{R(A)}$ satisfy the differential Dyson's equations

$$G_0^{-1}G^{R(A)} = 1 + \Sigma^{R(A)}G^{R(A)} \tag{2-25a}$$

$$G^{R(A)}G_0^{-1*} = 1 + G^{R(A)}\Sigma^{R(A)}, \tag{2-25b}$$

with $G_0^{-1}(11') = [\mathrm{i}\hbar\frac{\partial}{\partial t_1} - \hat{H}_0(\mathbf{r}_1) - U(1)]\delta(1-1')$ the (real-time) inverse of (2-12). Using (2-25a)-(2-25b) on (2-24) gives

$$G^{\lessgtr} = G^R G_0^{-1*}G_0^{\lessgtr}G_0^{-1}G^A + G^R\Sigma^{\lessgtr}G^A. \tag{2-26}$$

The first term in Eq. (2-26) corresponds to the contribution to the correlation functions $G^{\lessgtr}$ coming from the initial distribution of the system at time t_0 before the interactions and external field were turned on[4]. However, in steady-state we assume that a sufficiently long time has passed from the initial conditions in comparison with the relaxation time of the system (in fact, our initial choice of the Keldysh contour, in Fig. **2-1**, sets $t_0 \to -\infty$), and therefore this term will vanish out[5]. This last simplification leads to the so called *Keldysh equations* for the lesser and greater GFs

$$G^{\lessgtr}(11') = \int d2 \int d3 G^R(12)\Sigma^{\lessgtr}(23)G^A(31'). \tag{2-27}$$

2.2 Green's functions in energy space

In steady-state condition, the GFs and self-energies G^α, Σ^α ($\alpha \equiv R, A, <, >$) depend upon t and t' only via their difference (see Ref. [18], Sec. 24.2), that we will denote $\tau = t - t'$. We can then write $G^\alpha(\mathbf{r}, t; \mathbf{r}', t') \equiv G^\alpha(\mathbf{r}, \mathbf{r}'; \tau)$, which can be Fourier transformed to energy coordinates

$$\tilde{G}^\alpha(\mathbf{r}, \mathbf{r}'; E) = \int_{-\infty}^{\infty} d\tau\, e^{\frac{\mathrm{i}}{\hbar}E\tau}G^\alpha(\mathbf{r}, \mathbf{r}'; \tau) \tag{2-28}$$

$$G^\alpha(\mathbf{r}, \mathbf{r}'; \tau) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dE\, e^{-\frac{\mathrm{i}}{\hbar}E\tau}\tilde{G}^\alpha(\mathbf{r}, \mathbf{r}'; E). \tag{2-29}$$

[3] The reader can easily verify that $(1+G^R\Sigma^R)(1-G_0^R\Sigma^R) = 1$.

[4] In fact, since G_0^{-1} is a differential operator, we can further proceed integrating by parts and using the fact that $G_0^{-1}G_0^{\lessgtr} = G_0^{\lessgtr}G_0^{-1*} = 0$. In the end, one obtains an initial term of the form $\hbar^2 \int d\mathbf{r}_2 \int d\mathbf{r}_3 G^R(1; \mathbf{r}_2, t_0)G_0^{\lessgtr}(\mathbf{r}_2, t_0; \mathbf{r}_3, t_0)G^A(\mathbf{r}_3, t_0; 1')$. See Ref. [24].

[5] A more in-depth treatment of this can be found in Refs. [20, 25, 26]. See also [18] Sec. 26.6.3.

From (2-15), the retarded GF takes the following form in the energy domain

$$\begin{aligned}\tilde{G}^R(\mathbf{r},\mathbf{r}';E) &= \int_{-\infty}^{\infty} d\tau\, e^{\frac{\mathrm{i}}{\hbar}E\tau} G^R(\mathbf{r},\mathbf{r}';\tau) \\ &= -\mathrm{i}\int_{-\infty}^{\infty} d\tau\, e^{\frac{\mathrm{i}}{\hbar}E\tau}\theta(\tau)A(\mathbf{r},\mathbf{r}';\tau),\end{aligned} \tag{2-30}$$

where we have introduced the *spectral function*

$$A(\mathbf{r},\mathbf{r}';\tau) = \mathrm{i}[G^>(\mathbf{r},\mathbf{r}';\tau) - G^<(\mathbf{r},\mathbf{r}';\tau)] = \mathrm{i}[G^R(\mathbf{r},\mathbf{r}';\tau) - G^A(\mathbf{r},\mathbf{r}';\tau)]. \tag{2-31}$$

Applying the *convolution theorem* on (2-30)

$$\int_{-\infty}^{\infty} d\tau\, e^{\frac{\mathrm{i}}{\hbar}E\tau}\theta(\tau)A(\mathbf{r},\mathbf{r}';\tau) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dE'\,\tilde{\theta}(E-E')\tilde{A}(\mathbf{r},\mathbf{r}';E'), \tag{2-32}$$

and substituting the Fourier transform of the Heaviside step function

$$\tilde{\theta}(E) = \pi\hbar\,\delta(E) + \frac{\mathrm{i}\hbar}{E}, \tag{2-33}$$

it is finally obtained

$$\tilde{G}^R(\mathbf{r},\mathbf{r}';E) = -\frac{\mathrm{i}}{2}\tilde{A}(\mathbf{r},\mathbf{r}';E) + \mathcal{P}\int_{-\infty}^{\infty}\frac{dE'}{2\pi}\frac{\tilde{A}(\mathbf{r},\mathbf{r}';E')}{E-E'}, \tag{2-34}$$

where $\mathcal{P}$ denotes the *Cauchy principal value* of the integral. Eq. (2-34) directly relates the (Fourier transformed) spectral function $\tilde{A}$ with the imaginary part of the retarded GF $\tilde{G}^R$, and its real part with the *Hilbert transform* of $\tilde{A}$. A similar result can be obtained for the retarded self-energy

$$\tilde{\Sigma}^R(\mathbf{r},\mathbf{r}',E) = -\frac{\mathrm{i}}{2}\tilde{\Gamma}(\mathbf{r},\mathbf{r}',E) + \mathcal{P}\int_{-\infty}^{\infty}\frac{dE'}{2\pi}\frac{\tilde{\Gamma}(\mathbf{r},\mathbf{r}',E')}{E-E'}, \tag{2-35}$$

with the *broadening function* defined by

$$\Gamma(\mathbf{r},\mathbf{r}',\tau) = \mathrm{i}[\Sigma^>(\mathbf{r},\mathbf{r}',\tau) - \Sigma^<(\mathbf{r},\mathbf{r}',\tau)] = \mathrm{i}[\Sigma^R(\mathbf{r},\mathbf{r}',\tau) - \Sigma^A(\mathbf{r},\mathbf{r}',\tau)]. \tag{2-36}$$

In the following, in order to avoid unnecessarily complex notation, we shall drop the tilde when referring to Fourier transformed quantities $\tilde{G} \equiv G$, and will just refer to them with the same symbol of the time-dependent function. Eqs. (2-21) and (2-27) in the energy representation are given by

$$\begin{aligned}G^{R(A)}(\mathbf{r},\mathbf{r}',E) &= G_0^{R(A)}(\mathbf{r},\mathbf{r}',E) \\ &\quad + \int d\mathbf{r}_1\int d\mathbf{r}_2\, G_0^{R(A)}(\mathbf{r},\mathbf{r}_1,E)\Sigma^{R(A)}(\mathbf{r}_1,\mathbf{r}_2,E)G^{R(A)}(\mathbf{r}_2,\mathbf{r}',E)\end{aligned} \tag{2-37a}$$

$$G^{\lessgtr}(\mathbf{r},\mathbf{r}',E) = \int d\mathbf{r}_1\int d\mathbf{r}_2\, G^R(\mathbf{r},\mathbf{r}_1,E)\Sigma^{\lessgtr}(\mathbf{r}_1,\mathbf{r}_2,E)G^A(\mathbf{r}_2,\mathbf{r}',E), \tag{2-37b}$$

with the noninteracting GF given by

$$G_0^{R(A)}(\mathbf{r},\mathbf{r}',E) = [E + (-)\mathrm{i}\eta - \hat{H}_0(\mathbf{r}) - U(\mathbf{r})]^{-1}\delta(\mathbf{r}-\mathbf{r}'), \tag{2-38}$$

and the infinitesimal parameter $\eta \to 0^+$, which is necessary for the convergence of the Fourier transform. Symmetry relations (2-18a)-(2-18b) hold now in the form

$$[G^{\lessgtr}(\mathbf{r},\mathbf{r}',E)]^\dagger = -G^{\lessgtr}(\mathbf{r},\mathbf{r}',E) \tag{2-39a}$$

$$[G^{R}(\mathbf{r},\mathbf{r}',E)]^\dagger = G^{A}(\mathbf{r},\mathbf{r}',E). \tag{2-39b}$$

At thermal equilibrium, all GFs and self-energies are related by the *fluctuation-dissipation theorem* [18]

$$G^{<}(\mathbf{r},\mathbf{r}',E) = \mathrm{i}f_{FD}(E-E_F)A(\mathbf{r},\mathbf{r}',E) \tag{2-40}$$

$$G^{>}(\mathbf{r},\mathbf{r}',E) = \mathrm{i}[f_{FD}(E-E_F)-1]A(\mathbf{r},\mathbf{r}',E), \tag{2-41}$$

and

$$\Sigma^{<}(\mathbf{r},\mathbf{r}',E) = \mathrm{i}f_{FD}(E-E_F)\Gamma(\mathbf{r},\mathbf{r}',E) \tag{2-42}$$

$$\Sigma^{>}(\mathbf{r},\mathbf{r}',E) = \mathrm{i}[f_{FD}(E-E_F)-1]\Gamma(\mathbf{r},\mathbf{r}',E), \tag{2-43}$$

with f_{FD} the Fermi-Dirac distribution for the case of fermions, and E_F the equilibrium Fermi energy.

2.3 Physical observables

Once the GFs of a given system are known, it would be interesting to derive microscopic single particle properties as the mean value of some observable. For the purposes of carrier transport analysis in optoelectronic devices, we will be mainly interested in quantities as, *e.g.*, the electron and hole densities, the local density of states, the current density and the scattering rate associated to a particular type of interaction. Below we give a brief derivation of all these quantities in terms of the single particle GF of the many-particle system.

2.3.1 Carrier densities

Starting from the particle number operator in second quantization

$$\hat{n}(\mathbf{r},t) = \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r},t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t), \tag{2-44}$$

the average particle density at position $\mathbf{r}$ and time t is given by

$$\begin{aligned} n(\mathbf{r},t) &= \langle \hat{n}(\mathbf{r},t)\rangle \\ &= \langle \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r},t)\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t)\rangle \\ &= -\mathrm{i}\hbar \lim_{\mathbf{r}'\to\mathbf{r}} \lim_{t'\to t^+} G^{<}(\mathbf{r},t;\mathbf{r}',t'). \end{aligned} \tag{2-45}$$

In steady-state conditions, the GF can be Fourier transformed with respect to the time difference $\tau = t - t'$, yielding

$$\begin{aligned} n(\mathbf{r}) &= -\mathrm{i}\hbar \lim_{\mathbf{r}'\to\mathbf{r}} \lim_{t'\to t^+} G^<(\mathbf{r},\mathbf{r}';\tau) \\ &= -\mathrm{i}\hbar \lim_{\mathbf{r}'\to\mathbf{r}} \lim_{t'\to t^+} \int \frac{dE}{2\pi\hbar} e^{-\mathrm{i}\frac{E}{\hbar}\tau} G^<(\mathbf{r},\mathbf{r}';E) \\ &= -\mathrm{i} \int \frac{dE}{2\pi} G^<(\mathbf{r},\mathbf{r};E). \end{aligned} \tag{2-46}$$

A similar expression can be found for the holes, obtaining

$$p(\mathbf{r}) = \mathrm{i} \int \frac{dE}{2\pi} G^>(\mathbf{r},\mathbf{r};E). \tag{2-47}$$

By ignoring the energy integration in (2-46) and (2-47), we can define the energy resolved electron and hole densities as

$$n(\mathbf{r},E) = -\frac{\mathrm{i}}{2\pi} G^<(\mathbf{r},\mathbf{r};E) \tag{2-48a}$$

$$p(\mathbf{r},E) = \frac{\mathrm{i}}{2\pi} G^>(\mathbf{r},\mathbf{r};E). \tag{2-48b}$$

2.3.2 Local density of states

Using (2-40), the electron density (2-46) at equilibrium can be written as

$$n(\mathbf{r}) = \int \frac{dE}{2\pi} f_{\mathrm{FD}}(E - E_F) A(\mathbf{r},\mathbf{r};E), \tag{2-49}$$

hence defining the *local density of states* (LDOS) in the following way

$$D(\mathbf{r},E) = \frac{1}{2\pi} A(\mathbf{r},\mathbf{r};E), \tag{2-50}$$

where $A(\mathbf{r},\mathbf{r};E)$ is the Fourier transformed spectral function (2-31). From (2-48a)-(2-48b), we can see that the LDOS satisfy

$$D(\mathbf{r},E) = p(\mathbf{r},E) + n(\mathbf{r},E). \tag{2-51}$$

2.3.3 Current density

The average current density follows from the usual expression of the probability current in second quantization [20, 27]

$$\begin{aligned} \mathbf{J}(\mathbf{r},t) &= -\frac{\mathrm{i}\hbar}{2m_0} \langle \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r},t) \nabla_{\mathbf{r}} \hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) - \nabla_{\mathbf{r}} \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r},t) \hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \rangle \\ &= -\frac{\mathrm{i}\hbar}{2m_0} \lim_{\mathbf{r}'\to\mathbf{r}} [\nabla_{\mathbf{r}} - \nabla_{\mathbf{r}'}] \langle \hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}',t) \hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) \rangle \\ &= -\frac{\hbar^2}{2m_0} \lim_{t'\to t^+} \lim_{\mathbf{r}'\to\mathbf{r}} [\nabla_{\mathbf{r}} - \nabla_{\mathbf{r}'}] G^<(\mathbf{r},t;\mathbf{r}',t'). \end{aligned} \tag{2-52}$$

Recalling the velocity operator $\hat{\mathbf{v}} = \frac{1}{m_0}\hat{\mathbf{p}}$ and the noninteracting Hamiltonian [23]

$$\hat{H}_0(\mathbf{r}) = \frac{1}{2m_0}\hat{p}^2(\mathbf{r}) + \sum_i V(\mathbf{r} - \mathbf{R}_i), \tag{2-53}$$

with $\hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla_{\mathbf{r}}$ and $V(\mathbf{r} - \mathbf{R}_i)$ the potential energy due to the lattice atoms at positions $\mathbf{R}_i$, we obtain an expression for the current density

$$\mathbf{J}(\mathbf{r}, t) = \frac{\mathrm{i}\hbar}{2} \lim_{t'\to t^+} \lim_{\mathbf{r}'\to\mathbf{r}} [\hat{\mathbf{v}}(\mathbf{r}') - \hat{\mathbf{v}}(\mathbf{r})] G^<(\mathbf{r}, t; \mathbf{r}', t'), \tag{2-54}$$

and its divergence[6]

$$\nabla_{\mathbf{r}} \cdot \mathbf{J}(\mathbf{r}, t) = \lim_{t'\to t^+} \lim_{\mathbf{r}'\to\mathbf{r}} [\hat{H}_0(\mathbf{r}) - \hat{H}_0(\mathbf{r}')] G^<(\mathbf{r}, t; \mathbf{r}', t'). \tag{2-55}$$

A direct relation between the current divergence (2-55) and the single-particle self-energies can be found from the Dyson's equations, by subtracting (2-10a) and (2-10b) and taking the limit $\mathbf{r}' \to \mathbf{r}$, $t' \to t^+$, hence obtaining[7]

$$\begin{aligned}\lim_{1'\to 1^+} \left\{ \mathrm{i}\hbar \left(\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_{1'}} \right) G(11') \right\} - \lim_{1'\to 1^+} \left\{ [\hat{H}_0(\mathbf{r}_1) - \hat{H}_0(\mathbf{r}_{1'})] G(11') \right\} \\ = \lim_{1'\to 1^+} \left\{ \int d2 [\Sigma(12) G(21') - G(12)\Sigma(21')] \right\},\end{aligned} \tag{2-56}$$

where we have used the simplified notation $\lim_{1'\to 1^+} \equiv \lim_{t_{1'}\to t_1^+} \lim_{\mathbf{r}_{1'}\to\mathbf{r}_1}$. Taking the lesser component of (2-56) gives

$$\begin{aligned}-\frac{\partial n}{\partial t}(\mathbf{r}_1, t_1) - \nabla_{\mathbf{r}} \cdot \mathbf{J}(\mathbf{r}_1, t_1) = \lim_{1'\to 1^+} \int d2 [\Sigma^R(12) G^<(21') + \Sigma^<(12) G^A(21') \\ - G^R(12)\Sigma^<(21') - G^<(12)\Sigma^A(21')],\end{aligned} \tag{2-57}$$

with

$$\frac{\partial n}{\partial t}(\mathbf{r}_1, t_1) = -\mathrm{i}\hbar \lim_{1'\to 1^+} \left\{ \left(\frac{\partial}{\partial t_1} + \frac{\partial}{\partial t_{1'}} \right) G^<(11') \right\}, \tag{2-58}$$

the time derivative of the electron density (2-45), and we have replaced the current divergence from Eq. (2-55). Eq. (2-57) is the continuity equation, and the right hand side can be identified as the *net recombination rate*

$$\begin{aligned}R(\mathbf{r}_1, t_1) = \lim_{1'\to 1^+} \int d2 [\Sigma^R(12) G^<(21') + \Sigma^<(12) G^A(21') \\ - G^R(12)\Sigma^<(21') - G^<(12)\Sigma^A(21')].\end{aligned} \tag{2-59}$$

[6]This result is true because in the limit $\mathbf{r}' \to \mathbf{r}$, the potential energy terms in (2-53) cancel out, leaving just the kinetic energy term: $\lim_{\mathbf{r}'\to\mathbf{r}}[\hat{H}_0(\mathbf{r}') - \hat{H}_0(\mathbf{r})] = \lim_{\mathbf{r}'\to\mathbf{r}}[\frac{1}{2m_0}\hat{\mathbf{p}}^2(\mathbf{r}') - \frac{1}{2m_0}\hat{\mathbf{p}}^2(\mathbf{r})]$.

[7]Note that the term coming from the external potential cancels out in the limit, *i.e.*, $\lim_{1'\to 1^+}[U(1) - U(1')]G(11') = 0$.

In steady-sate conditions, the electron density (2-46) is time independent, and therefore its derivative vanishes. As a consequence, after Fourier transforming with respect to the time difference $\tau = t_1 - t_{1'}$, Eq. (2-57) becomes

$$\nabla_{\mathbf{r}} \cdot \mathbf{J}(\mathbf{r}) = -\frac{1}{2\pi\hbar}\int dE \int d\mathbf{r}' [\Sigma^R(\mathbf{r},\mathbf{r}';E)G^<(\mathbf{r}',\mathbf{r};E) + \Sigma^<(\mathbf{r},\mathbf{r}';E)G^A(\mathbf{r}',\mathbf{r};E) \\ - G^R(\mathbf{r},\mathbf{r}';E)\Sigma^<(\mathbf{r}',\mathbf{r};E) - G^<(\mathbf{r},\mathbf{r}';E)\Sigma^A(\mathbf{r}',\mathbf{r};E)]. \quad (2\text{-}60)$$

The right hand side of Eq. (2-60) should also vanish when integrated over the whole energy range, in order to guarantee the conservation of the total current, *i.e.*,

$$R(\mathbf{r}) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dE \int d\mathbf{r}' [\Sigma^R(\mathbf{r},\mathbf{r}';E)G^<(\mathbf{r}',\mathbf{r};E) + \Sigma^<(\mathbf{r},\mathbf{r}';E)G^A(\mathbf{r}',\mathbf{r};E) \\ - G^R(\mathbf{r},\mathbf{r}';E)\Sigma^<(\mathbf{r}',\mathbf{r};E) - G^<(\mathbf{r},\mathbf{r}';E)\Sigma^A(\mathbf{r}',\mathbf{r};E)] = 0. \quad (2\text{-}61)$$

When energy integration is restricted instead to the conduction (valence) band, Eq. (2-60) can be identified as the NEGF version of the semiclassical steady-state electron (hole) continuity equation, with the right hand side corresponding to the net electron (hole) recombination rate R_n (R_p). By ignoring the energy integration in Eq. (2-60), we can define the spectrally resolved current divergence

$$\nabla_{\mathbf{r}} \cdot \mathbf{J}(\mathbf{r},E) = -\frac{1}{2\pi\hbar}\int d\mathbf{r}' [\Sigma^R(\mathbf{r},\mathbf{r}';E)G^<(\mathbf{r}',\mathbf{r};E) + \Sigma^<(\mathbf{r},\mathbf{r}';E)G^A(\mathbf{r}',\mathbf{r};E) \\ - G^R(\mathbf{r},\mathbf{r}';E)\Sigma^<(\mathbf{r}',\mathbf{r};E) - G^<(\mathbf{r},\mathbf{r}';E)\Sigma^A(\mathbf{r}',\mathbf{r};E)]. \quad (2\text{-}62)$$

Since the self-energy $\Sigma = \sum_i \Sigma_i$ is composed of a sum of self-energy terms coming from different scattering mechanisms Σ_i, Eq. (2-62) can be used to compute the current divergence (recombination rate) due to an individual scattering mechanism i by inserting Σ_i in place of the total self-energy Σ. Integrated over the device volume, Eq. (2-62) gives, by means of the *Gauss theorem*, the (spectrally resolved) total current density leaving the boundaries of the interacting region, *i.e.*, the *global* transition rate [28]

$$\bar{R}(E) = -\int d\mathbf{S} \cdot \mathbf{J}(\mathbf{r},E) = -\int d\mathbf{r}\, \nabla_{\mathbf{r}} \cdot \mathbf{J}(\mathbf{r},E) \\ = \frac{1}{2\pi\hbar}\int d\mathbf{r} \int d\mathbf{r}' [\Sigma^R(\mathbf{r},\mathbf{r}';E)G^<(\mathbf{r}',\mathbf{r};E) + \Sigma^<(\mathbf{r},\mathbf{r}';E)G^A(\mathbf{r}',\mathbf{r};E) \\ - G^R(\mathbf{r},\mathbf{r}';E)\Sigma^<(\mathbf{r}',\mathbf{r};E) - G^<(\mathbf{r},\mathbf{r}';E)\Sigma^A(\mathbf{r}',\mathbf{r};E)]. \quad (2\text{-}63)$$

The double space integration in (2-63) allows to exchange factors in each term, which can

be rearranged and factored, yielding

$$\begin{aligned}\bar{R}(E) &= \frac{1}{2\pi\hbar}\int d\mathbf{r}\int d\mathbf{r}'\,\big\{\left[\Sigma^R(\mathbf{r},\mathbf{r}';E)-\Sigma^A(\mathbf{r},\mathbf{r}';E)\right]G^<(\mathbf{r}',\mathbf{r};E)\\ &\qquad -\Sigma^<(\mathbf{r},\mathbf{r}';E)\left[G^R(\mathbf{r}',\mathbf{r};E)-G^A(\mathbf{r}',\mathbf{r};E)\right]\big\}\\ &= \frac{1}{2\pi\hbar}\int d\mathbf{r}\int d\mathbf{r}'\,\big\{\left[\Sigma^R(\mathbf{r},\mathbf{r}';E)-\Sigma^A(\mathbf{r},\mathbf{r}';E)+\Sigma^<(\mathbf{r},\mathbf{r}';E)\right]G^<(\mathbf{r}',\mathbf{r};E)\\ &\qquad -\Sigma^<(\mathbf{r},\mathbf{r}';E)\left[G^R(\mathbf{r}',\mathbf{r};E)-G^A(\mathbf{r}',\mathbf{r};E)+G^<(\mathbf{r}',\mathbf{r};E)\right]\big\}\,,\end{aligned}$$

and using Eq. (2-17), gives the final result

$$\bar{R}(E) = \frac{1}{2\pi\hbar}\int d\mathbf{r}\int d\mathbf{r}'\,[\Sigma^>(\mathbf{r},\mathbf{r}';E)G^<(\mathbf{r}',\mathbf{r};E)-\Sigma^<(\mathbf{r},\mathbf{r}';E)G^>(\mathbf{r}',\mathbf{r};E)], \tag{2-64}$$

which is a simpler expression for the computation of the total scattering rate, with a very intuitive interpretation: $\Sigma^<(E)$ ($\Sigma^>(E)$) is proportional to the rate at which electrons are scattered into (out of) the state with energy E, while $G^<(E)$ ($G^>(E)$) is proportional to the number of electrons (vacancies or holes) occupying the state with energy E. Therefore, the product $\Sigma^<(E)G^>(E)$ represents the total *in-scattering* rate at energy E, and $\Sigma^>(E)G^<(E)$ represents the total *out-scattering* rate at energy E.

2.4 Spatial discretization

Given a suitable discrete basis $\{\phi_\nu(\mathbf{r})\}$, characterized by a set of quantum numbers ν, the GF can be discretized to numerically evaluate the Dyson's equation. This is done by expanding the field operators as linear combination of the given basis functions

$$\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t) = \sum_\nu \phi_\nu(\mathbf{r})\,\hat{c}_\nu(t) \tag{2-65a}$$

$$\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r},t) = \sum_\nu \phi^*_\nu(\mathbf{r})\,\hat{c}^\dagger_\nu(t), \tag{2-65b}$$

obtaining for the GF in Eq. (2-2)

$$\begin{aligned}G(\mathbf{r},t;\mathbf{r}',t') &= -\frac{\mathrm{i}}{\hbar}\langle\hat{T}_C\{\hat{\Psi}_{\mathcal{H}}(\mathbf{r},t)\hat{\Psi}^\dagger_{\mathcal{H}}(\mathbf{r}',t')\}\rangle\\ &= -\frac{\mathrm{i}}{\hbar}\langle\hat{T}_C\left\{\sum_\nu \phi_\nu(\mathbf{r})\hat{c}_\nu(t)\sum_\mu \phi^*_\mu(\mathbf{r}')\hat{c}^\dagger_\mu(t')\right\}\rangle\\ &= \sum_{\nu\mu}\phi_\nu(\mathbf{r})\left[-\frac{\mathrm{i}}{\hbar}\langle\hat{T}_C\{\hat{c}_\nu(t)\hat{c}^\dagger_\mu(t')\}\rangle\right]\phi^*_\mu(\mathbf{r}')\\ &= \sum_{\nu\mu}\phi_\nu(\mathbf{r})G_{\nu\mu}(t,t')\phi^*_\mu(\mathbf{r}'),\end{aligned} \tag{2-66}$$

or, in steady-state conditions

$$G(\mathbf{r},\mathbf{r}';E) = \sum_{\nu\mu} \phi_\nu(\mathbf{r}) G_{\nu\mu}(E) \phi^*_\mu(\mathbf{r}'). \tag{2-67}$$

The matrix $\mathbf{G}(E) = \{G_{\nu\mu}(E)\}$ in Eq. (2-67) is known as the *contravariant* representation of the GF. On the other hand, the *covariant* representation is given by the matrix $\tilde{\mathbf{G}}(E) = \{\tilde{G}_{\nu\mu}(E)\}$, such that

$$\tilde{G}_{\nu,\mu}(E) = \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_\nu(\mathbf{r}) G(\mathbf{r},\mathbf{r}';E) \phi_\mu(\mathbf{r}'). \tag{2-68}$$

Replacing (2-67) in (2-68), it is possible to find a relation between both representations

$$\begin{aligned}
\tilde{G}_{\nu,\mu}(E) &= \int d\mathbf{r} \int d\mathbf{r}' \phi^*_\nu(\mathbf{r}) \left[\sum_{\alpha\beta} \phi_\alpha(\mathbf{r}) G_{\alpha\beta}(E) \phi^*_\beta(\mathbf{r}') \right] \phi_\mu(\mathbf{r}') \\
&= \sum_{\alpha\beta} \left[\int d\mathbf{r} \phi^*_\nu(\mathbf{r}) \phi_\alpha(\mathbf{r}) \right] G_{\alpha\beta}(E) \left[\int d\mathbf{r}' \phi^*_\beta(\mathbf{r}') \phi_\mu(\mathbf{r}') \right] \\
&= \sum_{\alpha\beta} S_{\nu\alpha} G_{\alpha\beta}(E) S_{\beta\mu},
\end{aligned} \tag{2-69}$$

or, in matrix notation

$$\tilde{\mathbf{G}} = \mathbf{SGS}, \tag{2-70}$$

where we have introduced the *overlap matrix* $\mathbf{S} = \{S_{\nu\mu}\}$, given by the expression

$$S_{\nu\mu} = \int d\mathbf{r} \phi^*_\nu(\mathbf{r}) \phi_\mu(\mathbf{r}). \tag{2-71}$$

The same definitions are used to represent the self-energy $\Sigma(\mathbf{r},\mathbf{r}';E)$ in contravariant ($\boldsymbol{\Sigma}$) and covariant representation ($\tilde{\boldsymbol{\Sigma}}$). In the case of an orthonormal basis, $\mathbf{S}$ is the identity matrix and both *covariant* and *contravariant* representations are equivalent. The discretization of the Dyson's equation (2-37a) for the retarded and advanced GFs is obtained using (2-67), yielding

$$\begin{aligned}
\sum_{\nu\mu} \phi_\nu(\mathbf{r}) G^R_{\nu\mu}(E) \phi^*_\mu(\mathbf{r}') = &\sum_{\nu\mu} \phi_\nu(\mathbf{r}) G^R_{0,\nu\mu}(E) \phi^*_\mu(\mathbf{r}') \\
&+ \int d\mathbf{r}_2 \int d\mathbf{r}_3 \sum_{\nu_1\mu_1} \phi_{\nu_1}(\mathbf{r}) G^R_{0,\nu_1\mu_1}(E) \phi^*_{\mu_1}(\mathbf{r}_2) \\
&\times \sum_{\nu_2\mu_2} \phi_{\nu_2}(\mathbf{r}_2) \Sigma^R_{\nu_2\mu_2}(E) \phi^*_{\mu_2}(\mathbf{r}_3) \\
&\times \sum_{\nu_3\mu_3} \phi_{\nu_3}(\mathbf{r}_3) G^R_{\nu_3\mu_3}(E) \phi^*_{\mu_3}(\mathbf{r}')
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\nu\mu} \phi_\nu(\mathbf{r}) G^R_{0,\nu\mu}(E) \phi^*_\mu(\mathbf{r}') \\
&+ \sum_{\nu_1\mu_1} \sum_{\nu_2\mu_2} \sum_{\nu_3\mu_3} \phi_{\nu_1}(\mathbf{r}) G^R_{0,\nu_1\mu_1}(E) \Sigma^R_{\nu_2\mu_2}(E) G^R_{\nu_3\mu_3}(E) \phi^*_{\mu_3}(\mathbf{r}') \\
&\times \underbrace{\int d\mathbf{r}_2\, \phi^*_{\mu_1}(\mathbf{r}_2) \phi_{\nu_2}(\mathbf{r}_2)}_{S_{\mu_1\nu_2}} \underbrace{\int d\mathbf{r}_3\, \phi^*_{\mu_2}(\mathbf{r}_3) \phi_{\nu_3}(\mathbf{r}_3)}_{S_{\mu_2\nu_3}} \\
&= \sum_{\nu\mu} \phi_\nu(\mathbf{r}) G^R_{0,\nu\mu}(E) \phi^*_\mu(\mathbf{r}') + \sum_{\nu_1\mu_1} \sum_{\nu_3\mu_3} \phi_{\nu_1}(\mathbf{r}) G^R_{\nu_1\mu_1}(E) \\
&\times \underbrace{\left(\sum_{\nu_2\mu_2} S_{\mu_1\nu_2} \Sigma^R_{\nu_2\mu_2}(E) S_{\mu_2\nu_3} \right)}_{\tilde{\Sigma}_{\mu_1\nu_3}(E)} G^R_{\nu_3\mu_3}(E) \phi^*_{\mu_3}(\mathbf{r}') \\
&= \sum_{\nu\mu} \phi_\nu(\mathbf{r}) \left\{ \mathbf{G}^R_0(E) + \mathbf{G}^R_0(E) \tilde{\mathbf{\Sigma}}(E) \mathbf{G}^R(E) \right\}_{\nu\mu} \phi^*_\mu(\mathbf{r}'),
\end{aligned}
$$

which results in the matrix Dyson's equation

$$
\mathbf{G}^R(E) = \mathbf{G}^R_0(E) + \mathbf{G}^R_0(E) \tilde{\mathbf{\Sigma}}(E) \mathbf{G}^R(E), \tag{2-72}
$$

or equivalently

$$
\left[\left\{ \mathbf{G}^R_0(E) \right\}^{-1} - \tilde{\mathbf{\Sigma}}(E) \right] \mathbf{G}^R(E) = \mathbb{1}. \tag{2-73}
$$

The noninteracting GF is obtained from (2-38), by noting that

$$
\begin{aligned}
\int d\mathbf{r} \int d\mathbf{r}' \phi^*_\nu(\mathbf{r}) \left[E + \mathrm{i}\eta - \hat{H}_0(\mathbf{r}) - U(\mathbf{r}) \right] G^R_0(\mathbf{r}, \mathbf{r}', E) \phi_\mu(\mathbf{r}') &= \int d\mathbf{r} \int d\mathbf{r}' \phi^*_\nu(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \phi_\mu(\mathbf{r}') \\
&= \int d\mathbf{r} \phi^*_\nu(\mathbf{r}) \phi_\mu(\mathbf{r}) \\
&= S_{\nu\mu}.
\end{aligned}
$$

Further developing the left hand side gives

$$
\begin{aligned}
S_{\nu\mu} &= \int d\mathbf{r} \int d\mathbf{r}' \phi^*_\nu(\mathbf{r}) \left[E + \mathrm{i}\eta - \hat{H}_0(\mathbf{r}) - U(\mathbf{r}) \right] \sum_{\nu_1\mu_1} \phi_{\nu_1}(\mathbf{r}) G^R_{0,\nu_1\mu_1}(E) \phi^*_{\mu_1}(\mathbf{r}') \phi_\mu(\mathbf{r}') \\
&= \sum_{\nu_1\mu_1} \int d\mathbf{r} \phi^*_\nu(\mathbf{r}) \left[E + \mathrm{i}\eta - \hat{H}_0(\mathbf{r}) - U(\mathbf{r}) \right] \phi_{\nu_1}(\mathbf{r}) G^R_{0,\nu_1\mu_1}(E) \int d\mathbf{r}' \phi^*_{\mu_1}(\mathbf{r}') \phi_\mu(\mathbf{r}') \\
&= \sum_{\nu_1\mu_1} \left[(E + \mathrm{i}\eta) S_{\nu\nu_1} - H_{\nu\nu_1} - U_{\nu\nu_1} \right] G^R_{0,\nu_1\mu_1}(E) S_{\mu_1\mu},
\end{aligned}
$$

which in matrix notation reads

$$[(E+\mathrm{i}\eta)\mathbf{S}-\mathbf{H}-\mathbf{U}]\,\mathbf{G}_0^R(E)\mathbf{S}=\mathbf{S},$$

or equivalently

$$\mathbf{G}_0^R(E)=[(E+\mathrm{i}\eta)\mathbf{S}-\mathbf{H}-\mathbf{U}]^{-1}\,, \tag{2-74}$$

where $\eta\to 0^+$, and

$$H_{\nu\mu}=\int d\mathbf{r}\,\phi_\nu^*(\mathbf{r})\hat{H}_0(\mathbf{r})\phi_\mu(\mathbf{r}) \tag{2-75a}$$

$$U_{\nu\mu}=\int d\mathbf{r}\,\phi_\nu^*(\mathbf{r})U(\mathbf{r})\phi_\mu(\mathbf{r}). \tag{2-75b}$$

A similar procedure can be followed using Eq. (2-37b), obtaining the Keldysh equation in its discretized form

$$\mathbf{G}^{\lessgtr}(E)=\mathbf{G}^R(E)\tilde{\mathbf{\Sigma}}^{\lessgtr}(E)\mathbf{G}^A(E). \tag{2-76}$$

Hereafter, self-energies will always be expressed in the *covariant* representation, and to simplify notation, we will omit the tilde symbol ($\tilde{\Sigma}$ will be simply written as Σ). Conversely, the GFs will always be assumed to be in the *contravariant* representation.

2.5 Choice of basis: the $\mathbf{k}\cdot\mathbf{p}$ model

In order to properly model optoelectronic device operation, an accurate description of the band structure is needed, specially close to band edges where most of the important processes and interactions are concentrated. The $\mathbf{k}\cdot\mathbf{p}$ theory presents an efficient model, able to accurately describe the band structure of near high-symmetry points (*e.g.*, Γ, X, L) without the need of atomic resolution in the space discretization, and therefore making it suitable for large-scale device applications. We will specifically focus our attention on the analysis of layered structures, *i.e.*, semiconductor devices composed of stacked material layers along a given transport direction (which, for simplicity, will be assumed to be the z-axis), as illustrated in Fig. **2-2**. In the case of layered semiconductors, we follow the general ansatz used in the $\mathbf{k}\cdot\mathbf{p}$ *envelope function approach*[8] (EFA) [30], with our real-space basis written as [31]

$$\phi_{\nu\mathbf{k}}(\mathbf{r})=\frac{1}{\sqrt{A}}e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}t_i(z)u_{m0}(\mathbf{r}), \tag{2-77}$$

[8]Within the EFA the nanostructure wavefunction is assumed to follow the *ansatz* [29]

$$\Psi(\mathbf{r}_t,z)=\frac{1}{\sqrt{A}}e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}\sum_{m=1}^{N_b}u_{m0}(\mathbf{r})F_{m\mathbf{k}}(z).$$

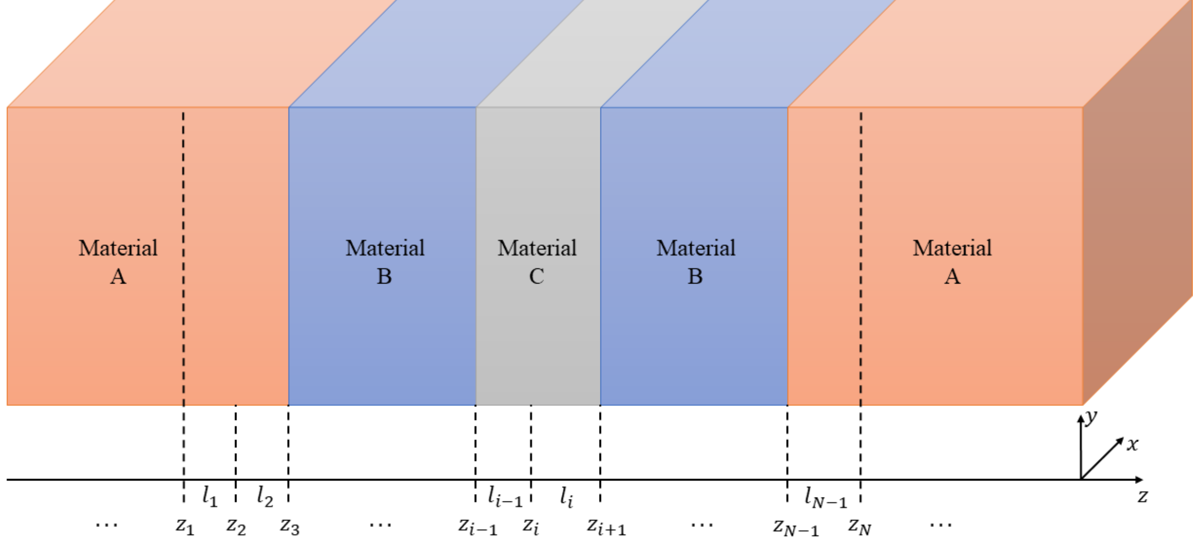


Figure 2-2: *Schematic diagram of a generic layered semiconductor device. Material layers are stacked along the z-axis, breaking the crystalline symmetry in this direction. Each material layer is assumed to be homogeneous along the transverse x-y plane. Within a finite-element implementation, the longitudinal 1-D space is partitioned into elements* $\{l_i, i \in \mathbb{Z}\}$ *by defining a mesh of discrete points* $\{z_i, i \in \mathbb{Z}\}$ *in the z-axis.*

where $\nu = (i, m)$ combines the space i and band m indices, $\mathbf{k} = (k_x, k_y)$ is the transversal component of the wavevector, $u_{m0}(\mathbf{r})$ is the zone-center lattice periodic Bloch function of band m, A is the cross sectional area for normalization and $\mathbf{r} = (\mathbf{r}_t, z)$ is position vector, with $\mathbf{r}_t$ the 2D component over the planes transversal to the z-axis. The $t_i(z)$ are shape functions which for our finite-element implementation will be peak-shaped functions centred at position z_i, defining in this way a discrete grid of points $\{z_i, i \in \mathbb{Z}\}$ in the z-axis. In $\mathbf{k}\cdot\mathbf{p}$ computations under the EFA the attention is restricted to the envelope function, which describes the behaviour of the nanostructure over the quantized direction, and therefore matrix elements are approximated using lattice-averaged quantities. For an arbitrary operator $\hat{O}$ acting mainly on the lattice periodic part, since the shape functions vary slowly over a unit cell Ω

The $\mathbf{k}$-dependent Bloch functions $u_{m\mathbf{k}}(\mathbf{r}) = \sum_{m=1}^{N_b} u_{m0}(\mathbf{r}) F_{m\mathbf{k}}(z)$ are expanded perturbatively as a sum of N_b zone-center periodic functions u_{m0}, with $F_{m\mathbf{k}}$ the slowly varying envelopes describing at every position z along the broken symmetry direction how the lattice-periodic functions are mixed together. Using a finite-element discretization, the envelope functions are approximated with Lagrange polynomials $F_{m\mathbf{k}}(z) = \sum_i F_{i,m}(\mathbf{k}) t_i(z)$, and thus the nanostructure wavefunction can be written as a linear combination of the basis functions (2-77), *i.e.*, $\Psi(\mathbf{r}) = \sum_i \sum_{m=1}^{N_b} F_{i,m}(\mathbf{k}) \phi_{im,\mathbf{k}}(\mathbf{r})$.

with respect to the lattice periodic Bloch functions, the matrix element can be approximated as [30]

$$\int d\mathbf{r}\, t_i(z)u^*_{m0}(\mathbf{r})\,\hat{O}\,t_j(z)u_{n0}(\mathbf{r}) \approx A\int dz\, t_i(z)t_j(z)\,\bar{O}(z), \tag{2-78}$$

with

$$\bar{O}(z) = \frac{1}{\Omega}\int_\Omega d\mathbf{r}'\, u^*_{m0}(\mathbf{r}')\,\hat{O}\,u_{n0}(\mathbf{r}'),$$

the (local) average of the operator over a unit cell volume Ω centred at z. From Eq. (2-78) the overlap matrix gives

$$\begin{aligned} S_{im,jn} &= \int d\mathbf{r}\,\phi^*_{im,\mathbf{k}}(\mathbf{r})\phi_{jn,\mathbf{k}}(\mathbf{r}) = \frac{1}{A}\int d\mathbf{r}\, t_i(z)t_j(z)\, u^*_{m0}(\mathbf{r})u_{n0}(\mathbf{r}) \\ &\approx \int dz\, t_i(z)t_j(z)\,\delta_{mn}, \end{aligned} \tag{2-79}$$

where the orthogonality condition of the zone-center Bloch functions was used

$$\frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r})u_{n0}(\mathbf{r}) = \delta_{mn}.$$

Note that due to our choice of shape functions, the overlap matrix $\mathbf{S}$ computed from (2-79) results to be block tri-diagonal. From (2-67) the steady-state GF expanded in our $\mathbf{k}\cdot\mathbf{p}$ basis reads[9]

$$G(\mathbf{r},\mathbf{r}';E) = \sum_{\mathbf{k}}\sum_{ijmn}\phi_{im,\mathbf{k}}(\mathbf{r})G_{im,jn}(\mathbf{k},E)\phi^*_{jn,\mathbf{k}}(\mathbf{r}'). \tag{2-80}$$

The physical observables can also be expressed in our discrete basis. Replacing (2-80) in (2-46), the electron density along the plane at position z_i gives

$$\begin{aligned} n(\mathbf{r}_t, z_i) &= -\mathrm{i}\int\frac{dE}{2\pi}G^<(\mathbf{r},\mathbf{r};E) \\ &= -\mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\sum_{i'j'mn}\phi_{i'm,\mathbf{k}}(\mathbf{r}_t,z_i)G^<_{i'm,j'n}(\mathbf{k},E)\phi^*_{j'n,\mathbf{k}}(\mathbf{r}_t,z_i) \\ &= -\mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\sum_{mn}u_{m0}(\mathbf{r}_t,z_i)G^<_{im,in}(\mathbf{k},E)u^*_{n0}(\mathbf{r}_t,z_i), \end{aligned} \tag{2-81}$$

[9] In general, the discrete form of the GF has a dual wavevector dependence, *i.e.*, $G(\mathbf{r},\mathbf{r}';E) = \sum_{\mathbf{k},\mathbf{k}'}\sum_{ijmn}\phi_{im,\mathbf{k}}(\mathbf{r})G_{im,jn}(\mathbf{k},\mathbf{k}',E)\phi^*_{jn,\mathbf{k}'}(\mathbf{r}')$, coming from both the basis functions. Under the assumption of homogeneity in the transverse direction, the GF $G(\mathbf{r},\mathbf{r}';E)$ depends only on the difference $\mathbf{r}_t - \mathbf{r}'_t$ and the discrete operator $G_{im,jn}(\mathbf{k},\mathbf{k}',E)$ results to be diagonal in $\mathbf{k}$ (see [32], Sec. 10.5). In (2-80) we therefore use the simplified notation $G_{im,jn}(\mathbf{k},\mathbf{k},E) \equiv G_{im,jn}(\mathbf{k},E)$.

where we have used that $t_j(z_i) = \delta_{ij}$. Since we are not interested in the atomic resolution provided by the Bloch functions but rather in averaged values, Eq. (2-81) can be further averaged over a crystal cell centred around z_i, obtaining[10]

$$\begin{aligned} n(z_i) &= -\mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\sum_{mn} G^{<}_{im,in}(\mathbf{k},E)\,\delta_{mn} \\ &= -\mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\mathrm{Tr}[\mathbf{G}^{<}_{i,i}(\mathbf{k},E)], \end{aligned} \tag{2-82}$$

where we have used the orthogonality property of functions u_{m0}, and the trace is over band indices. Similarly, the average hole density at z_i is written as

$$p(z_i) = \mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\mathrm{Tr}[\mathbf{G}^{>}_{i,i}(\mathbf{k},E)]. \tag{2-83}$$

From Eq. (2-51), the LDOS is related to the carrier densities by

$$\begin{aligned} D(z_i,E) &= n(z_i,E) + p(z_i,E) \\ &= \mathrm{i}\frac{1}{2\pi A}\sum_{\mathbf{k}}\mathrm{Tr}[\mathbf{G}^{>}_{i,i}(\mathbf{k},E) - \mathbf{G}^{<}_{i,i}(\mathbf{k},E)] \\ &= \frac{1}{2\pi A}\sum_{\mathbf{k}}\mathrm{Tr}[\mathbf{A}_{i,i}(\mathbf{k},E)], \end{aligned} \tag{2-84}$$

with the spectral function

$$\mathbf{A}(\mathbf{k},E) = \mathrm{i}[\mathbf{G}^{>}(\mathbf{k},E) - \mathbf{G}^{<}(\mathbf{k},E)]. \tag{2-85}$$

The computation of the discretized formula for the current can proceed following different approaches as shown in [27]. For simplicity, we simply show the result for a general non-orthogonal nearest-neighbour overlap basis. The particle current density flowing through the plane halfway between position z_i and z_{i+1} is then written

$$\begin{aligned} J_{i\rightarrow i+1} = \frac{1}{\hbar A}\sum_{\mathbf{k}}\int\frac{dE}{2\pi}\,\mathrm{Tr}\big[&(\mathbf{H}_{i,i+1}(\mathbf{k}) - E\mathbf{S}_{i,i+1})\mathbf{G}^{<}_{i+1,i}(\mathbf{k},E) \\ &- (\mathbf{H}_{i+1,i}(\mathbf{k}) - E\mathbf{S}_{i+1,i})\mathbf{G}^{<}_{i,i+1}(\mathbf{k},E)\big]. \end{aligned} \tag{2-86}$$

The components of the current coming from electron and hole transport can be separated by restricting the energy integral to the conduction or valence band, respectively. As stated before, energy resolved carrier and current densities can be defined by ignoring the energy

[10] The block matrix notation $\mathbf{G}_{i,j}$, in which the band indices are suppressed, will be commonly used from now as it is useful to simplify the notation in formulas involving matrix operations.

integration (2-82), (2-83) and (2-86). From Eq. (2-62), the average current divergence in its discretized form gives [31]

$$\frac{\partial J}{\partial z}(z_i, E) = -\frac{1}{2\pi\hbar A\Delta z_i}\sum_{\mathbf{k}} \mathrm{Tr}\Big[\big\{\mathbf{\Sigma}^R(\mathbf{k},E)\mathbf{G}^<(\mathbf{k},E) + \mathbf{\Sigma}^<(\mathbf{k},E)\mathbf{G}^A(\mathbf{k},E) - \mathbf{G}^R(\mathbf{k},E)\mathbf{\Sigma}^<(\mathbf{k},E) - \mathbf{G}^<(\mathbf{k},E)\mathbf{\Sigma}^A(\mathbf{k},E)\big\}_{i,i}\Big], \quad (2\text{-}87)$$

where the trace is over band indices of the i-th diagonal element resulting from the matrix product, and $\Delta z_i = 0.5(z_{i-1} - z_{i+1})$ is the mesh spacing. Current conservation requires (2-87) to vanish when integrated over the entire energy range.

2.5.1 The $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian for layered devices

Let us consider a layered semiconductor device as the one illustrated in Fig. **2-2**. A homogeneous crystalline structure is assumed along the transverse direction corresponding to the x-y plane. The discretized form of the noninteracting Hamiltonian can be obtained by applying (2-75a) to

$$\hat{H}_0(\mathbf{r}) = \frac{1}{2m_0}\hat{p}^2(\mathbf{r}) + V(\mathbf{r}) + \hat{H}_{so}(\mathbf{r}), \quad (2\text{-}88)$$

where $V(\mathbf{r})$ is the potential energy due to the lattice atoms[11], and we have introduced the spin-orbit energy

$$\hat{H}_{so}(\mathbf{r}) = \frac{1}{4m_0^2c^2}[\nabla V \times \hat{\mathbf{p}}]\cdot\sigma, \quad (2\text{-}89)$$

which up to this point had been omitted. The term σ is the Pauli spin matrix with components

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (2\text{-}90)$$

Applying $\hat{H}_0$ to the basis function $\phi_{jn,\mathbf{k}}$, and recalling that $\hat{\mathbf{p}} = -\mathrm{i}\hbar\nabla$, we obtain

$$\begin{aligned}
\hat{H}_0(\mathbf{r})\phi_{jn,\mathbf{k}}(\mathbf{r}) &= \left[\frac{1}{2m_0}\hat{p}^2(\mathbf{r}) + V(\mathbf{r}) + \hat{H}_{so}(\mathbf{r})\right]\phi_{jn,\mathbf{k}}(\mathbf{r}), \\
&= -\frac{\hbar^2}{2m_0\sqrt{A}}e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}\left[t_j(z)\nabla^2 u_{n0}(\mathbf{r}) + 2\mathrm{i}\underline{\mathbf{k}}\cdot\nabla u_{n0}(\mathbf{r})t_j(z) - \underline{k}^2 t_j(z)u_{n0}(\mathbf{r})\right] \\
&+ \frac{1}{\sqrt{A}}V(\mathbf{r})e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}t_j(z)u_{n0}(\mathbf{r}) + \frac{1}{4m_0^2c^2\sqrt{A}}e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}t_j(z)[\nabla V\times\hat{\mathbf{p}}]\cdot\sigma\, u_{n0}(\mathbf{r}) \\
&- \frac{\mathrm{i}\hbar}{4m_0^2c^2\sqrt{A}}[\nabla V\times\underline{\mathbf{k}}]\cdot\sigma\, e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}t_j(z)u_{n0}(\mathbf{r}),
\end{aligned} \quad (2\text{-}91)$$

[11] For simplicity, V is used here instead of the summation in Eq. (2-53).

where $\underline{\mathbf{k}} = (\mathbf{k}, k_z)$ is the three dimensional wavevector with $k_z \equiv -\mathrm{i}\frac{\partial}{\partial z}$. The last term in Eq. (2-91), which corresponds to a $\mathbf{k}$-dependent spin-orbit interaction, is usually neglected since the crystal momentum $\hbar\mathbf{k}$ is very small compared with the particle momentum $\mathbf{p}$, so most of the contribution to the spin-orbit energy comes from the fifth term on the right hand side. The noninteracting Hamiltonian matrix element (2-75a) is finally obtained by multiplying (2-91) by $\phi^*_{im,\mathbf{k}}$ on the left and integrating over space, obtaining

$$H_{im,jn} = \int dz\, t_i(z) t_j(z) \frac{1}{\Omega} \int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r}) \right] u_{n0}(\mathbf{r}) + \int dz\, t_i(z) H^{so}_{m,n} t_j(z)$$
$$+ \frac{\hbar}{m_0} \int dz\, t_i(z) \underline{\mathbf{k}} \cdot \mathbf{p}_{m,n} t_j(z) + \delta_{m,n} \frac{\hbar^2}{2m_0} \int dz\, t_i(z) \underline{k}^2 t_j(z), \tag{2-92}$$

where we have made use of the approximation (2-78) on every term, and

$$H^{so}_{m,n}(z) = \langle m|\hat{H}_{so}|n\rangle = \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) \hat{H}_{so}(\mathbf{r}) u_{n0}(\mathbf{r}), \tag{2-93}$$

$$\mathbf{p}_{m,n}(z) = \langle m|\hat{\mathbf{p}}|n\rangle = \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) \hat{\mathbf{p}}(\mathbf{r}) u_{n0}(\mathbf{r}). \tag{2-94}$$

are locally averaged quantities over a unit cell volume Ω centred at z. The lattice potential $V(\mathbf{r})$ is in general not periodic due to the inhomogeneities introduced along the z-axis. However, if the longitudinal integration is restricted to a single element $l_e = [z_e, z_{e+1}]$, then the potential energy $V(\mathbf{r})$ can be approximated to the periodic potential energy of a homogeneous crystal $V_p(\mathbf{r})$ made of the material present in the element, and we can use that the zone-center Bloch functions are the eigenfunctions of the spinless bulk Hamiltonian at $\mathbf{k} = 0$ [33]. The unit cell average in the first term of (2-92), inside a single element l_e, can be written as

$$\begin{aligned} \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r}) \right] u_{n0}(\mathbf{r}) &\approx \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) \left[-\frac{\hbar^2}{2m_0}\nabla^2 + V_p(\mathbf{r}) \right] u_{n0}(\mathbf{r}) \\ &= E_{n0} \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m0}(\mathbf{r}) u_{n0}(\mathbf{r}) \\ &= E_{n0}\, \delta_{m,n}, \end{aligned} \tag{2-95}$$

where $E_{n0} \equiv E_{n0}(z)$ is piecewise constant over each element, and so are $H^{so}_{m,n}(z)$ and $\mathbf{p}_{m,n}(z)$. The full integral over the longitudinal space in (2-92) can be broken down as a sum over element integrals

$$\int dz \to \sum_e \int_{l_e} dz,$$

and the last term on the right hand side of (2-92) can be expanded using integration by

parts[12]

$$\int dz\, t_i(z)\frac{\partial^2 t_j}{\partial z^2}(z) = t_i(z)\frac{\partial t_j}{\partial z}(z)\bigg|_{-\infty}^{\infty} - \int dz\, \frac{\partial t_i}{\partial z}(z)\frac{\partial t_j}{\partial z}(z). \tag{2-96}$$

The first term on the right hand side of Eq. (2-96) is equal to zero because the functions t_i are equal to zero far away from z_i, for any integer i. Using the previous considerations, the final result for the Hamiltonian matrix element is obtained

$$\begin{aligned} H_{im,jn} = \sum_e \Bigg[\left(E_{n0} + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{m,n} \int_{l_e} dz\, t_i(z) t_j(z) + H^{so}_{m,n} \int_{l_e} dz\, t_i(z) t_j(z) \\ + \frac{\hbar}{m_0} \int_{l_e} dz\, t_i(z)\, \underline{\mathbf{k}} \cdot \mathbf{p}_{m,n}\, t_j(z) + \delta_{m,n} \frac{\hbar^2}{2m_0} \int_{l_e} dz\, \frac{\partial t_i}{\partial z}(z) \frac{\partial t_j}{\partial z}(z) \Bigg], \end{aligned} \tag{2-97}$$

with

$$\int_{l_e} dz\, t_i(z)\, \underline{\mathbf{k}} \cdot \mathbf{p}\, t_j(z) = (k_x p_x + k_y p_y) \int_{l_e} dz\, t_i(z) t_j(z) - \mathrm{i} p_z \int_{l_e} dz\, t_i(z) \frac{\partial t_j}{\partial z}(z). \tag{2-98}$$

In practical implementations of Eq. (2-97), just a small number of bands is of interest (usually the bottom conduction bands and/or top valence bands), so the effect of coupling to the rest of the bands can be taken into account perturbatively, by means of Löwdin renormalization [33]. The main idea is to subdivide the complete set of bands into those of interest, which will be called class S, and the set of remote bands, which are called class R. For $n, m \in S$, the effective Hamiltonian accounting for the coupling to class R bands, to second order perturbation, has the following form [30]

$$H^{k\cdot p}_{im,jn} = H_{im,jn} + \sum_e \int_{l_e} dz\, t_i(z) \frac{\hbar^2}{m_0^2} \sum_{b\in B} \frac{(\underline{\mathbf{k}} \cdot \mathbf{p}_{m,b})(\underline{\mathbf{k}} \cdot \mathbf{p}_{b,n})}{E_{m0} - E_{b0}}\, t_j(z). \tag{2-99}$$

From Eqs. (2-97) and (2-99), it can be seen that the matrix element of the Hamiltonian in the finite-element $\mathbf{k}\cdot\mathbf{p}$ basis (2-77) is given by a second order polynomial in $\underline{\mathbf{k}}$. As in the case of the overlap matrix, the Hamiltonian $\mathbf{H}^{k\cdot p}$ is a block tri-diagonal matrix, since both the shape functions and their derivatives have nonzero overlap only when they correspond to adjacent points in the mesh.

Single band formulation

For the case of a single band, the Hamiltonian (2-99) reads (noting that $\mathbf{p}_{n,n} = 0$ and $H^{so}_{n,n} = 0$)

$$H^{k\cdot p}_{in,jn} = \sum_e \left[E_{n0} + \frac{\hbar^2}{2} \sum_{\alpha,\beta=x,y,z} \left(\frac{1}{m^*} \right)_{\alpha,\beta} k_\alpha k_\beta \right] \int_{l_e} dz\, t_i(z) t_j(z), \tag{2-100}$$

[12] Here the limits of integration extend from $-\infty$ to ∞ because up to this point we are considering the system composed of a device (the region of interest, where all the important interactions are concentrated) connected at both sides to semi-infinite particle reservoirs.

with $k_z \equiv -\mathrm{i}\frac{\partial}{\partial z}$, and the inverse effective mass tensor given by

$$\left(\frac{1}{m^*}\right)_{\alpha,\beta} = \frac{1}{m_0}\left(\delta_{\alpha,\beta} + \frac{1}{m_0}\sum_{n'\neq n}\frac{p^{\alpha}_{n,n'}p^{\beta}_{n',n} + p^{\beta}_{n,n'}p^{\alpha}_{n',n}}{E_{n0} - E_{n'0}}\right). \qquad \text{(2-101)}$$

When the effective mass tensor is diagonal, the Hamiltonian (2-100) is equivalent to the finite-element effective mass approximation with transversal and longitudinal masses m_t^* and m_l^*, respectively

$$H^{\mathrm{EMA}}_{i,j} = \sum_e \left[\left(E_0 + \frac{\hbar^2 k^2}{2m_t^*}\right)\int_{l_e} dz\, t_i(z)t_j(z) + \frac{\hbar^2}{2m_l^*}\int_{l_e} dz\,\frac{\partial t_i}{\partial z}(z)\frac{\partial t_j}{\partial z}(z)\right]. \qquad \text{(2-102)}$$

Multiband formulation

When multiple bands are considered, the Hamiltonian (2-99) is computed starting from the bulk $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian in the zone-center function basis including remote band perturbations, which can in general be written as a second order polynomial in $\underline{\mathbf{k}} = (k_x, k_y, k_z)$

$$\mathbf{H}^{k\cdot p}_{\mathrm{Bulk}}(\underline{\mathbf{k}}) = \sum_{\alpha,\beta=x,y,z} k_\alpha \bar{\mathbf{H}}^{(2)}_{\alpha,\beta} k_\beta + \sum_{\alpha=x,y,z} \bar{\mathbf{H}}^{(1)}_{\alpha} k_\alpha + \bar{\mathbf{H}}^{(0)}. \qquad \text{(2-103)}$$

For an N_b-band model, the matrices $\bar{\mathbf{H}}^{(i)}$, with $i = 0, 1, 2$, are $N_b \times N_b$ matrices whose values depend on crystal symmetry and on the zone-center functions used in the $\mathbf{k}\cdot\mathbf{p}$ representation[13]. Upon substitution of the longitudinal wavevector component by the corresponding differential operator $k_z \equiv -\mathrm{i}\frac{\partial}{\partial z}$, the Schrödinger equation for the nanostructure envelope function gives[14] [30, 29]

$$\left[\mathbf{H}^{(0)}(\mathbf{k},z) - \mathrm{i}\mathbf{H}^{(1)}_L(\mathbf{k},z)\frac{\partial}{\partial z} - \mathrm{i}\frac{\partial}{\partial z}\mathbf{H}^{(1)}_R(\mathbf{k},z) - \frac{\partial}{\partial z}\mathbf{H}^{(2)}(z)\frac{\partial}{\partial z}\right]\mathbf{F}_{\mathbf{k}}(z) = E\mathbf{F}_{\mathbf{k}}(z), \qquad \text{(2-104)}$$

with $\mathbf{F}_{\mathbf{k}}(z) = [F_{1\mathbf{k}}(z), ..., F_{N_b\mathbf{k}}(z)]^T$ the nanostructure envelope function, and the matrices $\mathbf{H}^{(i)}$, with $i = 0, 1, 2$, are formed by grouping terms that multiply differential operators of order i. The matrix $\mathbf{H}^{(1)}$ multiplying first order derivatives is split into left and right contributions depending on the position of the differential operator. This splitting is important

[13] Among the most important models we can find the Luttinger-Kohn model [34] for a 6-band Hamiltonian (heavy-hole, light-hole and spin-orbit split-off bands with their spin degeneracy), and the 8-band Hamiltonian which mixes the previous formulation with Kane's theory ([33], Ch. 4) to include the conduction band with its spin contribution. The 8-band Hamiltonian can be further block diagonalized into 4×4 diagonal submatrices [35] leading to the spinless 4-band formulation by considering just one block of the diagonal. Expressions for the bulk 8-band matrices can be found in App. A, both for the case of of Zinc-blende and Wurtzite crystals.

[14] The Schrödinger equation $\hat{H}_0\Psi = E\Psi$ is transformed into Eq. (2-104) for the envelope functions when the Hamiltonian is projected into the zone-center basis.

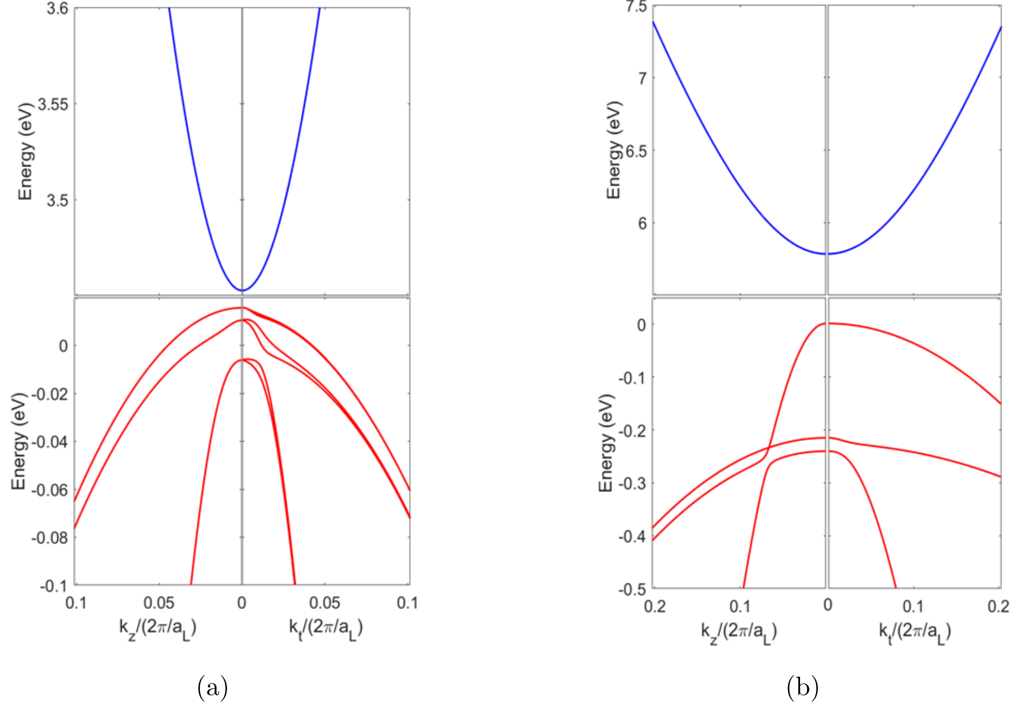


Figure 2-3: *Bulk band structure obtained from an 8-band* $\boldsymbol{k}\cdot\boldsymbol{p}$ *Hamiltonian for (a) GaN and (b) AlN. Band parameters can be found in Tab.* ***A-2****.*

as neglecting the particular order of the differential operators can lead to the appearance of spurious solutions in the numerical results (see [30] for further reading on operator ordering in $\mathbf{k}\cdot\mathbf{p}$ models with any number of confining directions). By applying Galerkin's procedure within a finite-element spatial discretization, the envelope functions are approximated using Lagrange polynomials $F_{n\mathbf{k}}(z) = \sum_j F_{j,n}(\mathbf{k})t_j(z)$, and Eq. (2-104) is multiplied by a test function $t_i(z)$ and integrated over space, leading to the weak formulation [36]

$$\begin{aligned}\sum_j \Bigg[&\int dz\, t_i(z)\mathbf{H}^{(0)}(\mathbf{k},z)t_j(z) \\ &-\mathrm{i}\int dz \left(t_i(z)\mathbf{H}_L^{(1)}(\mathbf{k},z)\frac{\partial t_j}{\partial z}(z) - \frac{\partial t_i}{\partial z}(z)\mathbf{H}_R^{(1)}(\mathbf{k},z)t_j(z)\right) \\ &+\int dz\, \frac{\partial t_i}{\partial z}(z)\mathbf{H}^{(2)}(z)\frac{\partial t_j}{\partial z}(z)\Bigg]\,\mathbf{F}_j(\mathbf{k}) = E\sum_j \int dz\, t_i(z)t_j(z)\,\mathbf{F}_j(\mathbf{k}),\end{aligned} \tag{2-105}$$

with $\mathbf{F}_j = [F_{j,1}, ..., F_{j,N_b}]^T$, defining in this way the multiband Hamiltonian matrix

$$\begin{aligned} \mathbf{H}^{k\cdot p}_{i,j}(\mathbf{k}) = \sum_e \Bigg[&\int_{l_e} dz\, t_i(z)\mathbf{H}^{(0)}(\mathbf{k},z)t_j(z) \\ &-\mathrm{i}\int_{l_e} dz \left(t_i(z)\mathbf{H}^{(1)}_L(\mathbf{k},z)\frac{\partial t_j}{\partial z}(z) - \frac{\partial t_i}{\partial z}(z)\mathbf{H}^{(1)}_R(\mathbf{k},z)t_j(z) \right) \\ &+\int_{l_e} dz\, \frac{\partial t_i}{\partial z}(z)\mathbf{H}^{(2)}(z)\frac{\partial t_j}{\partial z}(z) \Bigg], \end{aligned} \tag{2-106}$$

where we have expressed the spatial integral as a sum over element integrals. As described above, the matrices $\mathbf{H}^{(i)}$ are already known since they come from the bulk Hamiltonian, and are assumed to be piecewise constant over each element, as each element correspond to single bulk material. Expressions for the bulk $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian matrices can be found in App. A, with the corresponding band parameters. Fig. **2-3** shows the $\mathbf{k}\cdot\mathbf{p}$ band structure obtained for bulk GaN and AlN.

In writing (2-105) we have used integration by parts on the last two terms of the left hand side of (2-104), from which two boundary terms arise

$$\begin{aligned} -\mathrm{i}\int dz\, t_i(z)\frac{\partial}{\partial z}\left(\mathbf{H}^{(1)}_R(\mathbf{k},z)\mathbf{F}_\mathbf{k}(z)\right) = &-\mathrm{i}t_i(z)\mathbf{H}^{(1)}_R(\mathbf{k},z)\mathbf{F}_\mathbf{k}(z)\Big|_{z_L}^{z_R} \\ &+\mathrm{i}\int dz\, \frac{\partial t_i}{\partial z}(z)\mathbf{H}^{(1)}_R(\mathbf{k},z)\mathbf{F}_\mathbf{k}(z), \end{aligned} \tag{2-107}$$

and

$$\begin{aligned} -\int dz\, t_i(z)\frac{\partial}{\partial z}\left(\mathbf{H}^{(2)}(z)\frac{\partial \mathbf{F}_\mathbf{k}}{\partial z}(z)\right) = &-t_i(z)\mathbf{H}^{(2)}(z)\frac{\partial \mathbf{F}_\mathbf{k}}{\partial z}(z)\Big|_{z_L}^{z_R} \\ &+\int dz\, \frac{\partial t_i}{\partial z}(z)\mathbf{H}^{(2)}(z)\frac{\partial \mathbf{F}_\mathbf{k}}{\partial z}(z), \end{aligned} \tag{2-108}$$

with z_L and z_R the position of the left and right boundaries. Such boundary terms vanish when homogeneous Neumann and Dirichlet conditions are assumed (*i.e.*, when we consider a closed system) or when the integration limits are extended to infinity, as in (2-96). However, when the integration is restricted to the device region, open boundary conditions can be imposed by matching the boundary values of the envelope function and their derivatives with those of a semi-infinite reservoir at thermal equilibrium. Since the boundary values are in general complex, they introduce a nonzero imaginary part in the first and last block elements of the diagonal of the Hamiltonian, and therefore they represent a broadening of the confined levels as the Hamiltonian becomes non-Hermitian. This terms are strongly related with quantities referred as the boundary self-energies, which we will be treat in the next section.

Momentum matrix element

The momentum matrix element in our finite-element basis is given by

$$\begin{aligned}
\mathbf{\Pi}_{im,jn}(\mathbf{k}) &= \int d\mathbf{r}\, \phi^*_{im,\mathbf{k}}(\mathbf{r})\hat{\mathbf{p}}(\mathbf{r})\phi_{jn,\mathbf{k}}(\mathbf{r}) \\
&= \frac{1}{A}\int d\mathbf{r}\, t_i(z)u^*_{m0}(\mathbf{r})e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}\,\hat{\mathbf{p}}(\mathbf{r})\,t_j(z)u_{n0}(\mathbf{r})e^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t} \\
&= \frac{1}{A}\int d\mathbf{r}\, t_i(z)u^*_{m0}(\mathbf{r})(\hbar\mathbf{k}-\mathrm{i}\hbar\nabla)t_j(z)u_{n0}(\mathbf{r}) \\
&\approx \sum_e \left(\hbar\mathbf{k}\delta_{m,n}+\mathbf{p}_{m,n}\right)\int_{l_e} dz\, t_i(z)t_j(z).
\end{aligned} \tag{2-109}$$

with $\mathbf{p}_{m,n}$ given by (2-94). The momentum matrix element in the zone-center Bloch function basis $\mathbf{p}_{m,n}$ can be directly computed from the bulk $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian (2-103), by means of the *Feynman-Hellman* theorem

$$\mathbf{p} = \frac{m_0}{\hbar}\nabla_{\underline{\mathbf{k}}}\mathbf{H}^{k\cdot p}_{\text{Bulk}}. \tag{2-110}$$

In the band basis, the probability of optical transitions between different bands is proportional to the momentum matrix element [37, 35], and therefore the $\mathbf{k}$-dependent diagonal term is often neglected as it does not contribute to optical transitions

$$\mathbf{\Pi}_{im,jn} \approx \sum_e \mathbf{p}_{m,n}\int_{l_e} dz\, t_i(z)t_j(z). \tag{2-111}$$

2.6 Boundary conditions and boundary self-energy

As mentioned at the beginning of the chapter, we are considering a system composed of a finite region, referred as the *device*, in which the many-particle interactions are concentrated, connected to semi-infinite particle *reservoirs* or *contacts* at both sides, which are assumed to remain at thermal equilibrium, and which are characterized by Fermi energies μ_L and μ_R for the left and right contacts, respectively. Given the discrete mesh of points $\{z_i, i\in\mathbb{Z}\}$ defined by our finite-element basis, we will assume that the *device* region will correspond to the space between points z_1 and z_N. The left and right reservoir will consist of the mesh indices $\{-\infty,...,0\}$ and $\{N+1,...,\infty\}$, respectively. Defining $\tilde{\mathbf{H}} = E\mathbf{S}-\mathbf{H}-\mathbf{\Sigma}$, the Dyson's equation (2-73) in block notation, with each block representing one of the system regions, is given by

$$\begin{bmatrix} \tilde{\mathbf{H}}_L & \tilde{\mathbf{H}}_{LD} & \mathbf{0} \\ \tilde{\mathbf{H}}_{DL} & \tilde{\mathbf{H}}_D & \tilde{\mathbf{H}}_{DR} \\ \mathbf{0} & \tilde{\mathbf{H}}_{RD} & \tilde{\mathbf{H}}_R \end{bmatrix}
\begin{bmatrix} \mathbf{G}^R_L & \mathbf{G}^R_{LD} & \mathbf{G}^R_{LR} \\ \mathbf{G}^R_{DL} & \mathbf{G}^R_D & \mathbf{G}^R_{DR} \\ \mathbf{G}^R_{RL} & \mathbf{G}^R_{RD} & \mathbf{G}^R_R \end{bmatrix}
= \begin{bmatrix} \mathbb{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{1} \end{bmatrix}, \tag{2-112}$$

where $\tilde{\mathbf{H}}_{LD}$ and $\tilde{\mathbf{H}}_{DR}$ are the terms coupling device and reservoirs. In practical situations, it is necessary to avoid the inversion of an infinite dimension matrix for the computation of $\mathbf{G}^R$ in (2-112), but just to compute the submatrix corresponding to the device GF $\mathbf{G}_D^R$. In order to do this, we first define

$$\mathbf{g}_L^R = \tilde{\mathbf{H}}_L^{-1} = \left[E\mathbf{S}_L - \mathbf{H}_L - \mathbf{\Sigma}_L\right]^{-1} \tag{2-113a}$$
$$\mathbf{g}_R^R = \tilde{\mathbf{H}}_R^{-1} = \left[E\mathbf{S}_R - \mathbf{H}_R - \mathbf{\Sigma}_R\right]^{-1}, \tag{2-113b}$$

the GFs of the isolated left and right contacts, respectively, *i.e.*, the resulting GFs when the couplings $\tilde{\mathbf{H}}_{LD}$ and $\tilde{\mathbf{H}}_{DR}$ are set to zero. From (2-112), the following relations are obtained

$$\tilde{\mathbf{H}}_{DL}\mathbf{G}_{LD}^R + \tilde{\mathbf{H}}_D\mathbf{G}_D^R + \tilde{\mathbf{H}}_{DR}\mathbf{G}_{RD}^R = \mathbb{1} \tag{2-114a}$$
$$\left(\mathbf{g}_L^R\right)^{-1}\mathbf{G}_{LD}^R + \tilde{\mathbf{H}}_{LD}\mathbf{G}_D^R = \mathbf{0} \tag{2-114b}$$
$$\tilde{\mathbf{H}}_{RD}\mathbf{G}_D^R + \left(\mathbf{g}_R^R\right)^{-1}\mathbf{G}_{RD}^R = \mathbf{0}, \tag{2-114c}$$

where (2-113a) and (2-113b) have been used. Substituting $\mathbf{G}_{LD}^R$ and $\mathbf{G}_{RD}^R$ from (2-114b) and (2-114c) into (2-114a), leads to the final result

$$[\tilde{\mathbf{H}}_{RD} - \mathbf{\Sigma}^{RB}]\mathbf{G}_D^R = \mathbb{1}, \tag{2-115}$$

with

$$\mathbf{\Sigma}^{RB} = \tilde{\mathbf{H}}_{DL}\,\mathbf{g}_L^R\,\tilde{\mathbf{H}}_{LD} + \tilde{\mathbf{H}}_{DR}\,\mathbf{g}_R^R\,\tilde{\mathbf{H}}_{RD}, \tag{2-116}$$

the retarded boundary self-energy, with contributions coming from the left and right contacts. For the case of nearest-neighbour Hamiltonians, $\tilde{\mathbf{H}}_{LD}$ and $\tilde{\mathbf{H}}_{DR}$ are all zero matrices except for the bottom-left (band) block elements $(0,1)$ and $(N,N+1)$, respectively. Likewise, $\tilde{\mathbf{H}}_{DL}$ and $\tilde{\mathbf{H}}_{RD}$ are all zero matrices except for the upper-right (band) block elements $(1,0)$ and $(N+1,N)$, respectively. Therefore, the retarded boundary self-energy $\mathbf{\Sigma}^{RB}$ is a sparse matrix whose only nonzero $N_b \times N_b$ block elements are given by

$$\mathbf{\Sigma}_{1,1}^{RB} = \{\tilde{\mathbf{H}}_{DL}\}_{1,0}\{\mathbf{g}_L^R\}_{0,0}\{\tilde{\mathbf{H}}_{LD}\}_{0,1} \tag{2-117a}$$
$$\mathbf{\Sigma}_{N,N}^{RB} = \{\tilde{\mathbf{H}}_{DR}\}_{N,N+1}\{\mathbf{g}_R^R\}_{N+1,N+1}\{\tilde{\mathbf{H}}_{RD}\}_{N+1,N}. \tag{2-117b}$$

This means that in order to compute the boundary self-energy we do not need the whole information about the isolated contact GFs $\mathbf{g}_L^R$ and $\mathbf{g}_R^R$, but just their values in the immediate vicinity to the device. The lesser and greater components of the boundary self-energy can be computed using an analogous procedure to the one presented above, starting from the Keldysh equation (2-76), leading to the similar results

$$\mathbf{\Sigma}_{1,1}^{\lessgtr B} = \{\tilde{\mathbf{H}}_{DL}\}_{1,0}\{\mathbf{g}_L^{\lessgtr}\}_{0,0}\{\tilde{\mathbf{H}}_{LD}\}_{0,1} \tag{2-118a}$$
$$\mathbf{\Sigma}_{N,N}^{\lessgtr B} = \{\tilde{\mathbf{H}}_{DR}\}_{N,N+1}\{\mathbf{g}_R^{\lessgtr}\}_{N+1,N+1}\{\tilde{\mathbf{H}}_{RD}\}_{N+1,N}. \tag{2-118b}$$

Eqs. (2-117a)-(2-117b) and (2-118a)-(2-118b) are general results that hold even in the case when there are finite regions out of equilibrium, generally referred as the *leads*, that connect device and contact regions. In such cases, $\mathbf{g}_{L,R}$ represent the isolated GFs of the lead-contact system, and again the effect of the contacts on the leads is included through the boundary self-energy of the contacts (see Ref. [38]). When the *leads* are not included, the lesser and greater GFs of the isolated contacts follow the equilibrium relations (2-40) and (2-41), leading to the final results [23]

$$\mathbf{\Sigma}_{1,1}^{<B} = -f_{FD}(E-\mu_L)(\mathbf{\Sigma}_{1,1}^{RB} - \mathbf{\Sigma}_{1,1}^{AB}) \tag{2-119a}$$
$$\mathbf{\Sigma}_{N,N}^{<B} = -f_{FD}(E-\mu_R)(\mathbf{\Sigma}_{N,N}^{RB} - \mathbf{\Sigma}_{N,N}^{AB}), \tag{2-119b}$$

and

$$\mathbf{\Sigma}_{1,1}^{>B} = [1-f_{FD}(E-\mu_L)](\mathbf{\Sigma}_{1,1}^{RB} - \mathbf{\Sigma}_{1,1}^{AB}) \tag{2-120a}$$
$$\mathbf{\Sigma}_{N,N}^{>B} = [1-f_{FD}(E-\mu_R)](\mathbf{\Sigma}_{N,N}^{RB} - \mathbf{\Sigma}_{N,N}^{AB}), \tag{2-120b}$$

where $\mathbf{\Sigma}^{AB} = (\mathbf{\Sigma}^{RB})^\dagger$, f_{FD} is the Fermi-Dirac distribution, and $\mu_{L,R}$ are the Fermi energies of the left and right contacts, respectively.

2.6.1 Retarded boundary self-energy of the contacts

The general procedure followed in order to find the boundary self-energy $\mathbf{\Sigma}^{RB}$ is based in finding the surface contact GFs $\{\mathbf{g}_L^R\}_{0,0}$ and $\{\mathbf{g}_R^R\}_{N+1,N+1}$ by assuming a particular *ansatz* for the GFs based on the Bloch wave solutions of a semi-infinite bulk contact at equilibrium [38, 39]. In this work, we however show a different route more suitable for a finite-element implementation, by extracting the retarded boundary self-energy from the boundary terms resulting from the envelope Eq. (2-104) of the nanostructure as suggested in [40] and further generalized for a multiband implementation in [29]. Assuming that the integration boundaries in (2-107) and (2-108) are set as the boundaries of our *device* region, *i.e.*, $z_L = z_1$ and $z_R = z_N$, the boundary terms resulting from integration by parts are

$$-t_i(z)\mathbf{H}^{(2)}(z)\frac{\partial \mathbf{F_k}}{\partial z}(z)\bigg|_{z_1}^{z_N} - \mathrm{i}t_i(z)\mathbf{H}_R^{(1)}(\mathbf{k},z)\mathbf{F_k}(z)\bigg|_{z_1}^{z_N}. \tag{2-121}$$

Such terms vanish for every $i \notin \{1, N\}$, because for those values of i, any t_i evaluated at z_1 or z_N is equal to zero. We will first focus on the left boundary at z_1. For $i = 1$, the remaining terms are

$$\mathbf{H}^{(2)}(z_1)\frac{\partial \mathbf{F_k}}{\partial z}(z_1) + \mathrm{i}\mathbf{H}_R^{(1)}(\mathbf{k}, z_1)\mathbf{F_k}(z_1). \tag{2-122}$$

Following Ref. [39], we need to find the outgoing wave solution inside the semi-infinite bulk contacts[15], which for the left side can be written in general as a combination of left travelling (or decaying) Bloch waves [38], hence giving the following envelope solution at the contact side

$$\mathbf{F}_{\mathbf{k}}^{\text{out}}(z) = \sum_{m=1}^{N_b} a_m e^{-\mathrm{i}k_z^m z} \chi_m^L = \chi^L \mathbf{\Lambda} \mathbf{a}, \tag{2-123}$$

and its corresponding derivative

$$\frac{\partial \mathbf{F}_{\mathbf{k}}^{\text{out}}}{\partial z}(z) = -\mathrm{i} \sum_{m=1}^{N_b} a_m k_z^m e^{-\mathrm{i}k_z^m z} \chi_m^L = -\mathrm{i}\chi^L \mathbf{K}_z^L \mathbf{\Lambda} \mathbf{a} = -\mathrm{i}\chi^L \mathbf{K}_z^L [\chi^L]^{-1} \mathbf{F}_{\mathbf{k}}^{\text{out}}(z), \tag{2-124}$$

where $\mathbf{a} = [a_1, ..., a_{N_b}]^T$ are the Bloch wave expansion coefficients,

$$\mathbf{\Lambda} = \begin{pmatrix} e^{-\mathrm{i}k_z^1 z} & & 0 \\ & \ddots & \\ 0 & & e^{-\mathrm{i}k_z^{N_b} z} \end{pmatrix}, \tag{2-125}$$

is the diagonal matrix of propagation factors,

$$K_z^L = \begin{pmatrix} k_z^1 & & 0 \\ & \ddots & \\ 0 & & k_z^{N_b} \end{pmatrix}, \tag{2-126}$$

is the diagonal matrix of left propagating or decaying longitudinal wavevectors, and $\chi^L = [\chi_1^L, .., \chi_{N_b}^L]$ are the eigensolutions of the bulk Hamiltonian (2-103) corresponding to a left propagating or decaying longitudinal wavevector k_z^m at a fixed energy E. Using (2-124) in (2-122), and imposing the continuity of the envelope function and its derivative at z_1, leads to the boundary contribution to the Hamiltonian matrix element $\mathbf{H}_{1,1}^{k\cdot p}$ due to the coupling of the device to the left contact, which corresponds to the upper-left $N_b \times N_b$ block of the retarded boundary self-energy

$$\mathbf{\Sigma}_{1,1}^{RB}(\mathbf{k}, E) = -\mathrm{i}\mathbf{H}^{(2)}(z_1)\chi^L \mathbf{K}_z^L [\chi^L]^{-1} + \mathrm{i}\mathbf{H}_R^{(1)}(\mathbf{k}, z_1), \tag{2-127}$$

where the energy dependence is contained in the eigenmodes and eigenvectors, $\mathbf{K}_z^L$ and χ^L, respectively. Similarly, the right boundary contribution can be computed using $j = N$ in (2-121). Expressing now the outgoing wave solution as a combination of right travelling (or

[15] Just the outgoing solutions are of interest as any wave solution travelling to the device will correspond to a source term in Eq. (2-105) and will not enter directly in the Hamiltonian.

decaying) Bloch waves and following the same computations, the bottom-right $N_b \times N_b$ block of the retarded boundary self-energy results to be

$$\mathbf{\Sigma}^{RB}_{N,N}(\mathbf{k}, E) = -\mathrm{i}\mathbf{H}^{(2)}(z_N)\chi^R \mathbf{K}^R_z [\chi^R]^{-1} + \mathrm{i}\mathbf{H}^{(1)}_R(\mathbf{k}, z_N). \tag{2-128}$$

The task of computing the retarded boundary self-energy at this point is reduced to finding the outgoing (propagating or decaying) modes k_z^m and corresponding eigenvectors χ_m of the bulk contact, at fixed values of energy E and transversal crystal momentum $\mathbf{k}$, which is also known as a *complex band structure problem*. Recalling the decomposition (2-103), the bulk Hamiltonian of the (left or right) contact can be written as a second order polynomial in the longitudinal wavevector k_z

$$\mathbf{H} = \mathbf{H}^{(0)} + \mathbf{H}^{(1)} k_z + \mathbf{H}^{(2)} k_z^2, \tag{2-129}$$

with $\mathbf{H}^{(1)} = \mathbf{H}^{(1)}_L + \mathbf{H}^{(1)}_R$. The quadratic eigenvalue problem $(\mathbf{H} - E\mathbb{1})\chi = 0$, whose solution gives the states k_z and eigenvectors χ, can be transformed into the linear eigenvalue problem [41, 42]

$$\begin{pmatrix} 0 & \mathbb{1} \\ \mathbf{H}^{(0)} - E\mathbb{1} & \mathbf{H}^{(1)} \end{pmatrix} \begin{pmatrix} \chi \\ k_z\chi \end{pmatrix} = k_z \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbf{H}^{(2)} \end{pmatrix} \begin{pmatrix} \chi \\ k_z\chi \end{pmatrix}. \tag{2-130}$$

The complex eigenvalues occur in pairs, *i.e.*, propagating (decaying) states have the same real (imaginary) part and opposite signs. The $2N_b$ solutions of the complex band structure problem (2-130) are then classified according to their direction of motion (left or right in the longitudinal direction) and organized as the columns of matrices χ^L and χ^R, and the corresponding outgoing solutions are selected for each contact. The classification is done in the following way: propagating states (real k_z) are classified according to the sign of the group velocity v_g in the longitudinal direction, such that positive or negative velocities travel to the right or to the left, respectively. On the other hand, decaying states (complex k_z) are simply classified according to the sign of the imaginary part of the longitudinal wavevector, as decaying to the left ($\mathrm{Im}(k_z) < 0$) or decaying to the right ($\mathrm{Im}(k_z) > 0$). The computation of the group velocity can be carried out easily by means of the *Hellman-Feynman* theorem in the following way

$$v_g = \frac{1}{\hbar}\frac{dE}{dk_z}, \qquad \frac{dE}{dk_z} = \left\langle \chi^* \middle| \frac{d\mathbf{H}}{dk_z} \middle| \chi \right\rangle. \tag{2-131}$$

Once the retarded boundary self-energy is obtained, the lesser and greater self-energies follow directly from (2-119a)-(2-119b) and (2-120a)-(2-120b). For the single band case, whose Hamiltonian is given by Eq. (2-102), the retarded boundary self-energy from Eqs. (2-127) and (2-128) is straightforward

$$\mathbf{\Sigma}^{RB}_{1,1}(\mathbf{k}, E) = -\mathrm{i}\frac{\hbar^2}{2m_l^*} k_z^L, \qquad \mathbf{\Sigma}^{RB}_{N,N}(\mathbf{k}, E) = -\mathrm{i}\frac{\hbar^2}{2m_l^*} k_z^R, \tag{2-132}$$

where $k_z^{L,R}$ are obtained by inverting the energy dispersion relation $E = E_\mathrm{c}(z) + \hbar^2 k^2/2m_t^* + \hbar^2 k_z^2/2m_l^*$ at $z = z_1, z_N$, respectively, and E_c is the (conduction) band edge energy.

2.7 Particle interactions

As stated at the start of the chapter, particle interactions are included in the Dyson's equation (2-11a) by means of the scattering self-energy Σ, which can be obtained by expanding the GF in terms of powers of the interaction Hamiltonian $\hat{H}^i$ encoding all the different particle interactions. Starting from the definition of the contour-ordered GF (2-2), the nonequilibrium statistical average can be demonstrated to be equivalent to[16]

$$G(\mathbf{r},t;\mathbf{r}',t') = -\frac{\mathrm{i}}{\hbar}\frac{\langle \hat{T}_C\left\{\exp\left[-\frac{\mathrm{i}}{\hbar}\int_C d\tau \hat{H}^i_I(\tau)\right]\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\right\}\rangle_0}{\langle \hat{T}_{\tilde{C}}\left\{\exp\left[-\frac{\mathrm{i}}{\hbar}\int_C d\tau \hat{H}^i_I(\tau)\right]\right\}\rangle_0}, \tag{2-133}$$

where the field operators and interaction Hamiltonian are written in the interaction picture, denoted by the subscript I, and $\langle ... \rangle_0$ stands for an equilibrium ensemble average. The time ordered operator product in the numerator is defined as a power series of the perturbation Hamiltonian $\hat{H}^i_I$

$$\hat{T}_C\left\{\exp\left[-\frac{\mathrm{i}}{\hbar}\int_C d\tau \hat{H}^i_I(\tau)\right]\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\right\} = \\ \sum_{n=0}^{\infty}\left(\frac{-\mathrm{i}}{\hbar}\right)^n\frac{1}{n!}\int_C dt_1\cdots\int_C dt_n\,\hat{T}_C\{\hat{H}^i_I(t_1)\,\dots\,\hat{H}^i_I(t_n)\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\}, \tag{2-134}$$

which applied to (2-133) give for the zeroth-order and first-order terms (omitting the denominator)

$$G^{(0)}(\mathbf{r},t;\mathbf{r}',t') = -\frac{\mathrm{i}}{\hbar}\langle \hat{T}_C\left\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\right\}\rangle_0 = G_0(\mathbf{r},t;\mathbf{r}',t') \tag{2-135a}$$

$$G^{(1)}(\mathbf{r},t;\mathbf{r}',t') = \left(\frac{-\mathrm{i}}{\hbar}\right)^2\langle \hat{T}_C\left\{\int_C dt_1\,\hat{H}^i_I(t_1)\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\right\}\rangle_0. \tag{2-135b}$$

Eq. (2-135a) corresponds to the *free propagator*, *i.e.*, the GF of the unperturbed, noninteracting system (2-12). In order to proceed further with $G^{(1)}$, it is necessary to insert an expression for $\hat{H}^i_I$ depending on the specific type of interaction. Consider the *carrier-carrier interaction* in (2-4) which, written in the interaction picture, gives in second quantization

$$\hat{H}^i_I(t) = \frac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\,\hat{\Psi}^\dagger_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t)V(\mathbf{r}-\mathbf{r}')\hat{\Psi}_I(\mathbf{r}',t)\hat{\Psi}_I(\mathbf{r},t), \tag{2-136}$$

where the Coulomb potential is assumed to be instantaneous. Substituting in (2-135b) gives

$$G^{(1)}(\mathbf{r},t;\mathbf{r}',t') = \left(\frac{-\mathrm{i}}{\hbar}\right)^2\frac{1}{2}\int_C dt_1\int d\mathbf{r}_1\int d\mathbf{r}_2\,V(\mathbf{r}_1-\mathbf{r}_2) \\ \times\langle \hat{T}_C\left\{\hat{\Psi}^\dagger_I(\mathbf{r}_1,t_1)\hat{\Psi}^\dagger_I(\mathbf{r}_2,t_1)\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}^\dagger_I(\mathbf{r}',t')\right\}\rangle_0. \tag{2-137}$$

[16]We omit the derivation of this expression. The interested reader is referred to [43], Chs. 8 and 9.

We may now apply the *Wick-Matsubara theorem*, which states that the equilibrium statistical average of a time-ordered product of field operators is given by the sum of equilibrium statistical averages for all the possible pairs of field-operators [18, 43]

$$\langle \hat{T}\{\hat{O}_1\hat{O}_2...\hat{O}_m\}\rangle_0 = \sum_P (-1)^P \langle \hat{T}\{\hat{O}_1\hat{O}_2\}\rangle\langle \hat{T}\{\hat{O}_3\hat{O}_4\}\rangle...\langle \hat{T}\{\hat{O}_{m-1}\hat{O}_m\}\rangle, \tag{2-138}$$

where the sum is over all possible permutation of the m operators, and a minus one is added for each fermion field-operator permutation. Applying (2-138) to (2-137) gives

$$\begin{aligned}
&G^{(1)}(\mathbf{r},t;\mathbf{r}',t') = \left(\frac{-\mathrm{i}}{\hbar}\right)^2 \frac{1}{2}\int_C dt_1 \int d\mathbf{r}_1 \int d\mathbf{r}_2\, V(\mathbf{r}_1-\mathbf{r}_2)\\
&\times \big[\langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\\
&- \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\\
&+ \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\\
&- \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\\
&+ \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\\
&- \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r},t)\hat{\Psi}_I^\dagger(\mathbf{r}_1,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_2,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}_2,t_1)\}\rangle_0\, \langle \hat{T}_C\{\hat{\Psi}_I(\mathbf{r}_1,t_1)\hat{\Psi}_I^\dagger(\mathbf{r}',t')\}\rangle_0\big],
\end{aligned} \tag{2-139}$$

where we have omitted terms containing products of pairs of creation or annihilation operators, as their average is zero. Using (2-135a), is finally written as

$$\begin{aligned}
G^{(1)}(\mathbf{r},t;\mathbf{r}',t') = \mathrm{i}\hbar\frac{1}{2}\int_C dt_1 \int d\mathbf{r}_1 \int d\mathbf{r}_2\, V(\mathbf{r}_1-\mathbf{r}_2)&\\
\times \big[G_0(\mathbf{r}_1,t_1;\mathbf{r}_1,t_1)\, G_0(\mathbf{r}_2,t_1;\mathbf{r}_2,t_1)\, G_0(\mathbf{r},t;\mathbf{r}',t')&\\
- G_0(\mathbf{r}_1,t_1;\mathbf{r}_1,t_1)\, G_0(\mathbf{r},t;\mathbf{r}_2,t_1)\, G_0(\mathbf{r}_2,t_1;\mathbf{r}',t')&\\
+ G_0(\mathbf{r}_2,t_1;\mathbf{r}_1,t_1)\, G_0(\mathbf{r},t;\mathbf{r}_2,t_1)\, G_0(\mathbf{r}_1,t_1;\mathbf{r}',t')&\\
- G_0(\mathbf{r}_2,t_1;\mathbf{r}_1,t_1)\, G_0(\mathbf{r}_1,t_1;\mathbf{r}_2,t_1)\, G_0(\mathbf{r},t;\mathbf{r}',t')&\\
+ G_0(\mathbf{r},t,\mathbf{r}_1,t_1)\, G_0(\mathbf{r}_1,t_1;\mathbf{r}_2,t_1)\, G_0(\mathbf{r}_2,t_1;\mathbf{r}',t')&\\
- G_0(\mathbf{r},t;\mathbf{r}_1,t_1)\, G_0(\mathbf{r}_2,t_1;\mathbf{r}_2,t_1)\, G_0(\mathbf{r}_1,t_1;\mathbf{r}',t')\big].&
\end{aligned} \tag{2-140}$$

Eq. (2-140) presents the first order term of the series expansion of the full GF as a sum of terms containing the GF of the unperturbed system G_0. Higher order terms in the expansion series present a form similar to (2-140), *i.e.*, a power of the coupling term (in this case $V(\mathbf{r}_1 - \mathbf{r}_2)$) multiplied by a series of unperturbed GFs. The resulting terms can be conveniently expressed in terms of Feynman diagrams, for which the rules presented in Tab. **2-1** are used. It can be demonstrated that terms corresponding to disconnected diagrams are exactly cancelled by the denominator in (2-133) (see [21], Ch. 10). The resulting diagrammatic expansion of the GF, up to first order, is given by[17]

[17] Out of the six terms in (2-140), two correspond to disconnected diagrams and therefore they cancel out

Expression	Description	Diagram
$i\hbar G(\mathbf{r},t;\mathbf{r}',t')$	Full Green's function	$(\mathbf{r}',t')$ $(\mathbf{r},t)$
$i\hbar G_0(\mathbf{r},t;\mathbf{r}',t')$	Free Green's function	$(\mathbf{r}',t')$ $(\mathbf{r},t)$
$-iV(\mathbf{r}_1,\mathbf{r}_2)$	Coulomb interaction	$\mathbf{r}_1$ $\mathbf{r}_2$
$\int d\mathbf{r}_1 \int_C dt_1$	Intermediate variables	$(\mathbf{r}',t')$ $(\mathbf{r},t)$ $(\mathbf{r}_1,t_1)$
$(-1)\times$	Fermion loop	
$i\hbar D_0(\mathbf{r},t;\mathbf{r}',t')$	Free phonon	

Table 2-1: Definition of the equivalent of the Green's function in terms of Feynman diagrams. Continuous straight lines represent the unperturbed GF G_0 (*particle lines*); the arrow runs from the second argument to the first. The circle (Fermion loop) indicates the unperturbed GF G_0 with coincident initial and final points. The point where the interaction acts is called a *vertex*, where an integration is implied over intermediate variables (intermediate means excluding positions and times of the incoming and outgoing lines). Wavy lines connecting two vertices represent the interaction potential (*interaction lines*) and wiggly lines represent free phonons.

$(\mathbf{r}',t')$ $(\mathbf{r},t)$ $\approx$ + +

$=$ $(\mathbf{r}',t')$ $(\mathbf{r},t)$ + $(\mathbf{r}',t')$ $(\mathbf{r}_1,t_1)$ Σ_{HF} $(\mathbf{r}_2,t_2)$ $(\mathbf{r},t)$ (2-141)

or equivalently

$$G(\mathbf{r},t;\mathbf{r}',t') \approx G_0(\mathbf{r},t;\mathbf{r}',t') + \int d1 \int d2\, G_0(\mathbf{r},t;2)\Sigma_{HF}(2,1)G_0(1;\mathbf{r}',t'), \qquad (2\text{-}142)$$

with the denominator. The remaining four terms correspond to a double repetition of the Hartree and the Fock terms, hence cancelling the factor one half.

with

$$\Sigma_{HF} = \underbrace{\text{[Hartree diagram]}}_{\text{Hartree}} + \underbrace{\text{[Fock diagram]}}_{\text{Fock}}, \tag{2-143}$$

the Hartree-Fock self-energy. Including higher order terms in the expansion of the Green's function leads to a similar expression to (2-142), with the self-energy given by the sum of all possible combinations of *irreducible self-energy diagrams*[18]. By defining the *proper self-energy* Σ^* as the sum of all irreducible self-energies

$$\Sigma^* = \text{[Hartree]} + \text{[Fock]} + \underbrace{\text{[Direct]}}_{\text{Direct}} + \underbrace{\text{[Exchange]}}_{\text{Exchange}} + \ldots, \tag{2-144}$$

the total self-energy can be written as the sum of all possible repetitions of the proper self-energy

$$\Sigma = \Sigma^* + \Sigma^* \rightarrow \Sigma^* + \Sigma^* \rightarrow \Sigma^* \rightarrow \Sigma^* + \ldots . \tag{2-145}$$

Inserting Σ instead of Σ_{HF} in (2-142), gives for the full GF

$$\begin{aligned} (\mathbf{r}', t') \Rightarrow (\mathbf{r}, t) &= \longrightarrow + \longrightarrow \Sigma \longrightarrow \\ &= \longrightarrow + \longrightarrow \Sigma^* \\ &\times \left(\longrightarrow + \longrightarrow \Sigma^* \longrightarrow + \ldots \right) \\ &= (\mathbf{r}', t') \longrightarrow (\mathbf{r}, t) + (\mathbf{r}', t') \longrightarrow (\mathbf{r}_1, t_1)\, \Sigma^*\, (\mathbf{r}_2, t_2) \Rightarrow (\mathbf{r}, t) \end{aligned} \tag{2-146}$$

which is the diagrammatic version of the Dyson's equation (2-11a). The product in (2-146) represents an integral over intermediate variables. Eq. (2-144), where Σ^* is written as a

[18] The irreducible self-energy diagrams correspond to those diagrams that cannot be broken down into smaller irreducible diagrams by removing a G_0-line. See [21].

function of G_0, is known as the *Born approximation.* By collecting certain type of diagrams into a single irreducible diagram, Σ^* can be written as a function of the full GF. As an example, the Fock term reads

$$= \quad + \quad \Sigma^* \quad + \quad \ldots$$

$$= \quad + \quad + \quad + \quad + \quad + \quad \ldots . \tag{2-147}$$

The resulting expression for the self-energy[19]

$$\Sigma^* \quad = \quad + \quad + \quad + \quad + \quad \ldots , \tag{2-148}$$

Direct Exchange

referred as the *self-consistent Born approximation*, defines a self-consistent procedure for the determination of the GF, as it both follows from and determines the self-energy Σ^*. In practical implementations, the self-energy is approximated with few diagrams from (2-148), which in our particular case will be the Fock term for both phonon and photon scattering, and the Hartree-Fock plus direct collision terms for electron-electron scattering (also known as the GW-approximation). From now on we will be always treating with the proper self-energy Σ^*, so we will drop the asterisk and refer to it simply as the self-energy.

2.7.1 Electron-phonon self-energy

The interaction of electrons with phonons is described by the electron-phonon potential (see [27] and [44], Sec. 1.3)

$$V_{ep}(\mathbf{r},t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{q}} U_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} (\hat{a}_{\mathbf{q}}(t) + \hat{a}^{\dagger}_{-\mathbf{q}}(t)), \tag{2-149}$$

where $\hat{a}^{\dagger}_{\mathbf{q}}$ and $\hat{a}_{\mathbf{q}}$ are the creation and annihilation operators for a phonon in mode $\mathbf{q} = (\mathbf{q}_t, q_z)$ lying in the first Brillouin zone, $U_{\mathbf{q}}$ is a term containing the Fourier transform of the electron-ion potential, and V is the normalization volume. The full derivation of the electron-boson

[19] The four diagrams shown in the right hand side of Eq. (2-148), *i.e.*, Hartree-Fock, direct-collision and exchange-collision, are all the diagrams up to second order in interaction, in the SCBA. The approximation obtained using only this four terms is known as *second Born approximation.*

self-energy is given in the App. B for the interested reader. Comparing Eqs. (2-149) and (B-1), it follows that the interaction matrix element (B-4) is given by

$$M_{im,jn}(\mathbf{k},\mathbf{k}',\mathbf{q}) = \frac{1}{\sqrt{V}}U_{\mathbf{q}}\int d\mathbf{r}\,\phi^*_{im,\mathbf{k}}(\mathbf{r})e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\phi_{jn,\mathbf{k}'}(\mathbf{r}), \tag{2-150}$$

where we can replace the basis functions $\phi^*_{im,\mathbf{k}}$ and $\phi_{jn,\mathbf{k}}$ using Eq. (2-77), obtaining

$$\begin{aligned} M_{im,jn}(\mathbf{k},\mathbf{k}',\mathbf{q}) &= \frac{1}{\sqrt{V}}U_{\mathbf{q}}\frac{1}{A}\int d\mathbf{r}\,t_i(z)u^*_{m0}(\mathbf{r})e^{\mathrm{i}(\mathbf{q}_t-\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_t}e^{\mathrm{i}q_z z}t_j(z)u_{n0}(\mathbf{r}) \\ &\approx \frac{1}{\sqrt{V}}U_{\mathbf{q}}\frac{1}{A}\int d\mathbf{r}\,t_i(z)e^{\mathrm{i}q_z z}t_j(z)e^{\mathrm{i}(\mathbf{q}_t-\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_t}\,\frac{1}{\Omega}\int_\Omega u^*_{m0}(\mathbf{r})u_{n0}(\mathbf{r}) \\ &= \frac{1}{\sqrt{V}}U_{\mathbf{q}}\int dz\,t_i(z)e^{\mathrm{i}q_z z}t_j(z)\frac{1}{A}\int d\mathbf{r}_t\,e^{\mathrm{i}(\mathbf{q}_t-\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_t}\,\delta_{m,n} \\ &= \frac{1}{\sqrt{V}}U_{\mathbf{q}}\int dz\,t_i(z)e^{\mathrm{i}q_z z}t_j(z)\,\delta_{\mathbf{k}-\mathbf{q}_t,\mathbf{k}'}\,\delta_{m,n}\,, \end{aligned} \tag{2-151}$$

with $\mathbf{r}_t$ and $\mathbf{q}_t$ are the transversal components of the position and phonon momentum vectors, respectively, and we have used the approximation (2-78). Inserting (2-151) in the Fock terms of the electron-phonon self-energy (B-35) and (B-36), and assuming a long-wavelength approximation [31]

$$\int dz\,t_i(z)e^{\mathrm{i}q_z z}t_j(z)\delta_{m,n} \approx e^{\mathrm{i}q_z z_i}S_{im,jn} \approx e^{\mathrm{i}q_z z_j}S_{im,jn},$$

leads to the final expression, in block matrix notation

$$\begin{aligned} \mathbf{\Sigma}^{\lessgtr}_{i,j}(\mathbf{k},E) &= \frac{1}{V}\sum_{\mathbf{q}}|U_{\mathbf{q}}|^2\,e^{\mathrm{i}q_z(z_i-z_j)} \\ &\quad\times\left\{\mathbf{S}\big[(n_{\mathbf{q}}+1)\mathbf{G}^{\lessgtr}(\mathbf{k}-\mathbf{q}_t,E\pm\hbar\omega_{\mathbf{q}})+n_{\mathbf{q}}\mathbf{G}^{\lessgtr}(\mathbf{k}-\mathbf{q}_t,E\mp\hbar\omega_{\mathbf{q}})\big]\mathbf{S}\right\}_{i,j} \end{aligned} \tag{2-152}$$

$$\begin{aligned} \mathbf{\Sigma}^{R}_{i,j}(\mathbf{k},E) &= \frac{1}{V}\sum_{\mathbf{q}}|U_{\mathbf{q}}|^2\,e^{\mathrm{i}q_z(z_i-z_j)} \\ &\quad\times\Big\{\mathbf{S}\Big[\frac{1}{2}\big(\mathbf{G}^{<}(\mathbf{k}-\mathbf{q}_t,E-\hbar\omega_{\mathbf{q}})-\mathbf{G}^{<}(\mathbf{k}-\mathbf{q}_t,E+\hbar\omega_{\mathbf{q}})\big) \\ &\quad+(n_{\mathbf{q}}+1)\mathbf{G}^{R}(\mathbf{k}-\mathbf{q}_t,E+\hbar\omega_{\mathbf{q}})+n_{\mathbf{q}}\mathbf{G}^{R}(\mathbf{k}-\mathbf{q}_t,E-\hbar\omega_{\mathbf{q}})\Big]\mathbf{S}\Big\}_{i,j}, \end{aligned} \tag{2-153}$$

where $n_{\mathbf{q}} = \left(e^{\beta\hbar\omega_{\mathbf{q}}}-1\right)^{-1}$ is the phonon occupation number, given by the Bose-Einstein distribution of a phonon with angular frequency $\omega_{\mathbf{q}}$ and mode $\mathbf{q}$, and $\beta^{-1} = k_B T$. Expressions (2-152) and (2-153) assume an equilibrium population of bulk phonons modes, *i.e.*, renormalization of the phonon population due to electron-phonon interaction is neglected. In writing (2-153) we have also neglected the principal value integral from the retarded self-energy (B-36), as it does not lead to relaxation or dephasing but just to a small energy

renormalization that is assumed to be negligible [27, 23, 31]. An alternative expression for the retarded self-energy can be obtained by using Eq. (2-35). Since the lesser (greater) self-energies can be interpreted as in(out)-scattering rates, Eq. (2-152) presents a very intuitive interpretation: two terms contribute to the in-scattering at an energy E for the lesser self-energy, the first is proportional to $n_{\mathbf{q}}+1$ and represents both the processes of stimulated and spontaneous emission of phonons through the relaxation of electrons at an energy $E+\hbar\omega_{\mathbf{q}}$. The second term, which is proportional to $n_{\mathbf{q}}$, describe the absorption of phonons through the excitation of electrons at an energy $E-\hbar\omega_{\mathbf{q}}$. A similar interpretation can be obtained from the greater self-energy in terms of holes.

Acoustic phonons

For acoustic phonons, a linear phonon dispersion $\omega_{\mathbf{q}} = u_l q$ is assumed, and the scattering strength is given by [44, 45]

$$U_{\mathbf{q}} = \sqrt{\frac{\hbar D_a^2}{2\rho u_l} q} \tag{2-154}$$

where u_l is the longitudinal sound velocity in the material, D_a is the acoustic deformation potential, $V = AL$ is the normalization volume and ρ is the material (mass) density. At sufficiently high temperature, $\hbar\omega_{\mathbf{q}}$ is much smaller than $k_B T$, allowing to approximate the phonon occupation number $n_{\mathbf{q}} \approx \frac{k_B T}{\hbar\omega_{\mathbf{q}}} \approx n_{\mathbf{q}}+1$. Using this result and Eqs. (2-152) and (2-153), the acoustic phonon self-energy yields

$$\boldsymbol{\Sigma}_{i,j}^{\lessgtr,R}(\mathbf{k},E) = \frac{1}{\Delta_{ij}} \frac{D_a^2 k_B T}{A\rho u_l^2} \sum_{\mathbf{q}_t} \left\{ \mathbf{S}\mathbf{G}^{\lessgtr,R}(\mathbf{k}-\mathbf{q}_t,E)\mathbf{S} \right\}_{i,j}, \tag{2-155}$$

with[20]

$$\frac{1}{\Delta_{ij}} = \frac{1}{L}\sum_{q_z} e^{\mathrm{i}q_z(z_i-z_j)} \approx \frac{1}{2\pi}\int_{-\pi/a_L}^{\pi/a_L} dq_z\, e^{\mathrm{i}q_z(z_i-z_j)} = \frac{1}{\pi}\frac{\sin\frac{\pi}{a_L}(z_i-z_j)}{z_i-z_j}, \tag{2-156}$$

and a_L the lattice constant in the transport direction. Note that in (2-155) we have used an elastic approximation, *i.e.*, $E \pm \hbar\omega_{\mathbf{q}} \approx E$ and therefore the resulting self-energy does not involve energy relaxation.

[20] Throughout this work, a common approximation used to compute sums over the first Brillouin zone in the reciprocal space will be

$$\sum_{\mathbf{q}} \approx \frac{V}{(2\pi)^3}\int_{FBZ} d^3\mathbf{q}.$$

Polar optical phonons

Interaction with polar optical phonons, according to the Fröhlich theory of polar optical scattering, assumes a dispersion-less longitudinal phonon energy $\hbar\omega_{\mathrm{LO}}$, and a scattering strength term given by [44, 45]

$$U_{\mathbf{q}} = \sqrt{\frac{e^2\hbar\omega_{\mathrm{LO}}}{2}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}\right)}\frac{q}{q^2+q_0^2}, \tag{2-157}$$

where ϵ_s and ϵ_∞ are the static and optical dielectric constants of the material, respectively, and $q_0 = \sqrt{\frac{e^2 n}{\epsilon_s k_B T}}$ is the inverse Debye-Hückel screening length for a screening medium with carrier density n. Substituting the scattering strength (2-157) in the self-energy expressions (2-152) and (2-153) gives

$$\begin{aligned}\mathbf{\Sigma}^{\lessgtr}_{i,j}(\mathbf{k},E) = {} & \frac{e^2\hbar\omega_{\mathrm{LO}}}{2V}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}\right)\sum_{\mathbf{q}}\left(\frac{q}{q^2+q_0^2}\right)^2 e^{\mathrm{i}q_z(z_i-z_j)} \\ & \times \left\{\mathbf{S}\left[(n_{\mathrm{LO}}+1)\mathbf{G}^{\lessgtr}(\mathbf{k}-\mathbf{q}_t, E\pm\hbar\omega_{\mathrm{LO}}) + n_{\mathrm{LO}}\mathbf{G}^{\lessgtr}(\mathbf{k}-\mathbf{q}_t, E\mp\hbar\omega_{\mathrm{LO}})\right]\mathbf{S}\right\}_{i,j}\end{aligned} \tag{2-158}$$

$$\begin{aligned}\mathbf{\Sigma}^{R}_{i,j}(\mathbf{k},E) = {} & \frac{e^2\hbar\omega_{\mathrm{LO}}}{2V}\left(\frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}\right)\sum_{\mathbf{q}}\left(\frac{q}{q^2+q_0^2}\right)^2 e^{\mathrm{i}q_z(z_i-z_j)} \\ & \times \Big\{\mathbf{S}\Big[\frac{1}{2}\big(\mathbf{G}^{<}(\mathbf{k}-\mathbf{q}_t, E-\hbar\omega_{\mathrm{LO}}) - \mathbf{G}^{<}(\mathbf{k}-\mathbf{q}_t, E+\hbar\omega_{\mathrm{LO}})\big) \\ & + (n_{\mathrm{LO}}+1)\mathbf{G}^{R}(\mathbf{k}-\mathbf{q}_t, E+\hbar\omega_{\mathrm{LO}}) + n_{\mathrm{LO}}\mathbf{G}^{R}(\mathbf{k}-\mathbf{q}_t, E-\hbar\omega_{\mathrm{LO}})\Big]\mathbf{S}\Big\}_{i,j},\end{aligned} \tag{2-159}$$

with the longitudinal optical phonon occupation number given by

$$n_{\mathrm{LO}} = \left(e^{\beta\hbar\omega_{\mathrm{LO}}} - 1\right)^{-1}, \tag{2-160}$$

and $\beta^{-1} = k_B T$. At this point, the substitution $\mathbf{q}'_t = \mathbf{k} - \mathbf{q}_t$ is made, as it simplifies the numerical computation of the self-energy (evaluation of the GF at $\mathbf{k}-\mathbf{q}_t$ could be difficult as such point could lie outside of the k-space grid chosen for the numerical computation). Although this substitution would imply a displacement of the summation region (from the Brillouin zone to the same region displaced by $\mathbf{k}$), in practical situations just small values of $\mathbf{k}$ are needed as the GFs become negligible far away from the Γ point (the GFs are computed up a cut-off value k_{max}), so we will neglect this displacement. The summation can be approximated by an integral in the following way

$$\begin{aligned}& \frac{1}{V}\sum_{\mathbf{q}}\left(\frac{q}{q^2+q_0^2}\right)^2 e^{\mathrm{i}q_z(z_i-z_j)} = \frac{1}{V}\sum_{\mathbf{q}'}\left(\frac{\sqrt{|\mathbf{k}-\mathbf{q}'_t|^2+q_z^2}}{|\mathbf{k}-\mathbf{q}'_t|^2+q_z^2+q_0^2}\right)^2 e^{\mathrm{i}q_z(z_i-z_j)} \\ & \approx \frac{1}{(2\pi)^3}\int dq'_t\, q'_t \int_{-\frac{\pi}{a_L}}^{\frac{\pi}{a_L}} dq_z\, e^{\mathrm{i}q_z(z_i-z_j)} \int_0^{2\pi} d\theta\, \frac{k^2+q'^2_t+q_z^2-2kq'_t\cos(\theta)}{(k^2+q'^2_t+q_z^2+q_0^2-2kq'_t\cos(\theta))^2}.\end{aligned} \tag{2-161}$$

In the *axial approximation* (isotropy for the transverse dispersion), the GFs are independent of θ, so the integral over the azimuthal angle can be performed directly in (2-161). In order to compute the integral, the identity $2\cos(\theta) = e^{i\theta} + e^{-i\theta}$ and the change of variable $z = e^{i\theta}$ should be used, transforming the angular integral into an integral over the unitary circle in the complex plane, and the residue theorem can then be applied (the explicit computations are omitted here, but they can be found in [23], Sec. 4.3.2), leading to the final result

$$\begin{aligned}\mathbf{\Sigma}^{\lessgtr}_{i,j}(\mathbf{k},E) &= \frac{e^2\hbar\omega_{\mathrm{LO}}}{4\pi^2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)\int dq_t\, q_t\, F(q_t,s_{ij},k) \\ &\quad\times\left\{\mathbf{S}\left[(n_{\mathrm{LO}}+1)\mathbf{G}^{\lessgtr}(\mathbf{q}_t,E\pm\hbar\omega_{\mathrm{LO}})+n_{\mathrm{LO}}\mathbf{G}^{\lessgtr}(\mathbf{q}_t,E\mp\hbar\omega_{\mathrm{LO}})\right]\mathbf{S}\right\}_{i,j} \qquad (2\text{-}162)\end{aligned}$$

$$\begin{aligned}\mathbf{\Sigma}^{R}_{i,j}(\mathbf{k},E) &= \frac{e^2\hbar\omega_{\mathrm{LO}}}{4\pi^2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)\int dq_t\, q_t\, F(q_t,s_{ij},k) \\ &\quad\times\Big\{\mathbf{S}\Big[\frac{1}{2}\big(\mathbf{G}^{<}(\mathbf{q}_t,E-\hbar\omega_{\mathrm{LO}})-\mathbf{G}^{<}(\mathbf{q}_t,E+\hbar\omega_{\mathrm{LO}})\big) \\ &\quad+(n_{\mathrm{LO}}+1)\mathbf{G}^{R}(\mathbf{q}_t,E+\hbar\omega_{\mathrm{LO}})+n_{\mathrm{LO}}\mathbf{G}^{R}(\mathbf{q}_t,E-\hbar\omega_{\mathrm{LO}})\Big]\mathbf{S}\Big\}_{i,j}, \qquad (2\text{-}163)\end{aligned}$$

with $s_{ij} = z_i - z_j$ and

$$\begin{aligned}F(q,s_{ij},k) &= \int_0^{\pi/a_L} dq_z \cos(q_z s_{ij}) \\ &\times\left[\frac{1}{\sqrt{(k^2+q^2+q_z^2+q_0^2)^2-4k^2q^2}} - q_0^2\frac{k^2+q^2+q_z^2+q_0^2}{((k^2+q^2+q_z^2+q_0^2)^2-4k^2q^2)^{3/2}}\right]. \qquad (2\text{-}164)\end{aligned}$$

with a_L the lattice parameter of the material. In Eq. (2-164) just the cosine term of the complex exponential has been preserved, because the term proportional to sine gives zero when integrated due to the symmetry of the sine function.

2.7.2 Electron-photon self-energy

The electron-photon potential, describing electromagnetic interaction of light waves with electrons, is given by [46]

$$V_{e\gamma}(\mathbf{r},t) = \frac{e}{m_0}\hat{\mathbf{A}}(\mathbf{r},t)\cdot\hat{\mathbf{p}}(\mathbf{r}), \qquad (2\text{-}165)$$

with the electromagnetic vector potential $\hat{\mathbf{A}}(\mathbf{r},t)$, in second quantization, given by

$$\hat{\mathbf{A}}(\mathbf{r},t) = \sum_{\lambda,\mathbf{q}} \mathbf{A}_0(\lambda,\mathbf{q})\, e^{i\mathbf{q}\cdot\mathbf{r}}(\hat{b}_{\lambda,\mathbf{q}}(t)+\hat{b}^\dagger_{\lambda,-\mathbf{q}}(t)), \qquad (2\text{-}166)$$

where $\hat{b}^\dagger_{\mathbf{q}}$ and $\hat{b}_{\mathbf{q}}$ are the creation and annihilation operators for a photon in mode λ and wavevector $\mathbf{q}$, the quantity

$$\mathbf{A}_0(\lambda,\mathbf{q}) = \sqrt{\frac{\hbar}{2V\epsilon_0\omega_{\mathbf{q}}}}\,\mathbf{e}_{\lambda,\mathbf{q}} \qquad (2\text{-}167)$$

and $\mathbf{e}_{\lambda,\mathbf{q}}$ the polarization vector of the electromagnetic field. The full derivation of the electron-photon self-energy (as in the case of the electron-phonon self-energy) is given in the App. B. A direct comparison of Eqs. (2-165) and (B-1) leads to the following form of the interaction matrix element (B-4)

$$M_{im,jn}(\mathbf{k},\mathbf{k}',\mathbf{q},\lambda) = \frac{e}{m_0}\,\mathbf{A}_0(\lambda,\mathbf{q})\cdot\int d\mathbf{r}\,\phi^*_{im,\mathbf{k}}(\mathbf{r})e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\,\hat{\mathbf{p}}(\mathbf{r})\,\phi_{jn,\mathbf{k}'}(\mathbf{r}), \tag{2-168}$$

which after substituting the basis functions from Eq. (2-77) gives

$$\begin{aligned}
M_{im,jn}(\mathbf{k},\mathbf{k}',\mathbf{q},\lambda) &= \frac{e}{m_0}\,\mathbf{A}_0(\lambda,\mathbf{q})\cdot\frac{1}{A}\int d\mathbf{r}\,t_i(z)u^*_{m0}(\mathbf{r})e^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_t}e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\,\hat{\mathbf{p}}(\mathbf{r})\,e^{\mathrm{i}\mathbf{k}'\cdot\mathbf{r}_t}t_j(z)u_{n0}(\mathbf{r})\\
&= \frac{e}{m_0}\,\mathbf{A}_0(\lambda,\mathbf{q})\cdot\frac{1}{A}\int d\mathbf{r}\,t_i(z)u^*_{m0}(\mathbf{r})e^{\mathrm{i}(\mathbf{q}_t-\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_t}e^{\mathrm{i}q_z z}\,(\hbar\mathbf{k}'-\mathrm{i}\hbar\nabla)t_j(z)u_{n0}(\mathbf{r})\\
&\approx \frac{e}{m_0}\,\mathbf{A}_0(\lambda,\mathbf{q})\cdot\int dz\,t_i(z)\,e^{\mathrm{i}q_z z}\,t_j(z)\,\frac{1}{A}\int d\mathbf{r}_t\,e^{\mathrm{i}(\mathbf{q}_t-\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}_t}\\
&\times\frac{1}{\Omega}\int_\Omega d\mathbf{r}\,u^*_{m0}(\mathbf{r})(\hbar\mathbf{k}'-\mathrm{i}\hbar\nabla)u_{n0}(\mathbf{r})\\
&\approx \frac{e}{m_0}\,\mathbf{A}_0(\lambda,\mathbf{q})\cdot\int dz\,t_i(z)\,e^{\mathrm{i}q_z z}\,t_j(z)\,\delta_{\mathbf{k}-\mathbf{q}_t,\mathbf{k}'}\,(\hbar\mathbf{k}'\delta_{m,n}+\mathbf{p}_{m,n}).
\end{aligned} \tag{2-169}$$

If the structure length is considerably shorter than the wavelength of the incident illumination, the *dipole approximation* can be applied by taking $e^{\mathrm{i}q_z z}\approx 1$, resulting in the following expression for the interaction matrix element

$$M_{im,jn}(\mathbf{k},\mathbf{k}',\mathbf{q},\lambda) \approx \delta_{\mathbf{k}-\mathbf{q}_t,\mathbf{k}'}\,\frac{e}{m_0}\sqrt{\frac{\hbar}{2V\epsilon_0\omega_\mathbf{q}}}\,\mathbf{e}_{\lambda,\mathbf{q}}\cdot\mathbf{\Pi}_{im,jn}(\mathbf{k}'), \tag{2-170}$$

where $\mathbf{\Pi}_{im,jn}$ is the momentum matrix element from Eq. (2-111). Since in practical numerical implementations, photon wavevectors are usually much smaller than the mesh spacing in k-space (for a monochromatic source with $\hbar\omega_\mathbf{q} = 1$ eV, we have $q \approx 5\ \mu\mathrm{m}^{-1}$ which is smaller than a typical mesh space of $\Delta k = 20\ \mu\mathrm{m}^{-1}$) [31], the approximation $\mathbf{k}-\mathbf{q}_t \approx \mathbf{k}$ is also assumed. Inserting (2-170) in the self-energy expressions (B-35) and (B-36), gives the final expressions for the electron-photon Fock self-energies, which read in full matrix notation

$$\begin{aligned}
\mathbf{\Sigma}^{\lessgtr}(\mathbf{k},E) &= \frac{e^2\hbar^2}{2V\epsilon_0 m_0^2}\sum_{\mathbf{q},\lambda}\frac{1}{\hbar\omega_\mathbf{q}}\\
&\times(\mathbf{e}_{\lambda,\mathbf{q}}\cdot\mathbf{\Pi})\left[(n_{\lambda,\mathbf{q}}+1)\,G^{\lessgtr}(\mathbf{k},E\pm\hbar\omega_\mathbf{q}) + n_{\lambda,\mathbf{q}}\,G^{\lessgtr}(\mathbf{k},E\mp\hbar\omega_\mathbf{q})\right](\mathbf{e}_{\lambda,\mathbf{q}}\cdot\mathbf{\Pi})
\end{aligned} \tag{2-171}$$

$$\begin{aligned}
\mathbf{\Sigma}^{R}(\mathbf{k},E) &= \frac{e^2\hbar^2}{2V\epsilon_0 m_0^2}\sum_{\mathbf{q},\lambda}\frac{1}{\hbar\omega_\mathbf{q}}\,(\mathbf{e}_{\lambda,\mathbf{q}}\cdot\mathbf{\Pi})\left[\frac{1}{2}\big(\mathbf{G}^<(\mathbf{k},E-\hbar\omega_\mathbf{q}) - \mathbf{G}^<(\mathbf{k},E+\hbar\omega_\mathbf{q})\big)\right.\\
&\left. + (n_{\lambda,\mathbf{q}}+1)\mathbf{G}^R(\mathbf{k},E+\hbar\omega_\mathbf{q}) + n_{\lambda,\mathbf{q}}\mathbf{G}^R(\mathbf{k},E-\hbar\omega_\mathbf{q})\right](\mathbf{e}_{\lambda,\mathbf{q}}\cdot\mathbf{\Pi})\,.
\end{aligned} \tag{2-172}$$

The sum over $\mathbf{q}$ is usually approximated as an integral over the incoming photon energy by using the identity $E_\gamma = \hbar\omega_{\mathbf{q}} = \hbar c\mathbf{q}$, with c the speed of light, such that

$$
\begin{aligned}
\frac{1}{V}\sum_q &\approx \frac{1}{(2\pi)^3}\int dq\, q^2 \int_0^{2\pi} d\theta \int_0^{\pi} d\phi \sin(\phi) \\
&= \frac{1}{(2\pi\hbar c)^3}\int dE_\gamma\, E_\gamma^2 \int d\Omega,
\end{aligned}
$$

where $d\Omega = d\theta\, d\phi \sin(\phi)$ is an infinitesimal solid angle element, and the angular integral is done over all possible directions of the incoming illumination. For the particular case of a *monochromatic single-mode source* of light, emitting at a frequency ω_γ, the expressions of the lesser and greater self-energies become

$$
\begin{aligned}
\boldsymbol{\Sigma}^{\lessgtr}(\mathbf{k}, E) = {} & \frac{e^2\hbar\, n_\gamma}{2V\epsilon_0 m_0^2\omega_\gamma}\left(\mathbf{e}_\gamma\cdot\boldsymbol{\Pi}\right)\left[G^{\lessgtr}(\mathbf{k}, E\pm\hbar\omega_\gamma) + G^{\lessgtr}(\mathbf{k}, E\mp\hbar\omega_\gamma)\right]\left(\mathbf{e}_\gamma\cdot\boldsymbol{\Pi}\right) \\
& + \frac{e^2}{4\pi^2\epsilon_0\hbar c^3 m_0^2}\sum_\lambda\int dE'\, E'\left(\mathbf{e}_\lambda\cdot\boldsymbol{\Pi}\right) G^{\lessgtr}(\mathbf{k}, E\pm E')\left(\mathbf{e}_\lambda\cdot\boldsymbol{\Pi}\right),
\end{aligned}
\tag{2-173}
$$

where $\mathbf{e}_\gamma$ is the polarization vector of the incoming illumination, and the photon number n_γ is related to the incoming light intensity I_γ (in units W/m^2) through the following relation

$$
n_\gamma = \frac{I_\gamma V\sqrt{\mu_r\epsilon_r}}{\hbar\omega_\gamma c},
\tag{2-174}
$$

with ϵ_r and μ_r the relative dielectric constant and magnetic permeability of the material, respectively. Note that although a single light mode is involved in both the absorption and stimulated emission processes, spontaneous emission can occur at any energy and therefore the integral over photon energies must be maintained in (2-173) (second term in the right hand side). The retarded self-energy for the case of monochromatic illumination can be found from the lesser and greater self-energies through Eq. (2-35).

2.7.3 Carrier-carrier self-energy

The main derivations concerning the perturbation expansion of the carrier-carrier self-energy were partially discussed in Sec. 2.7, leading diagrammatic expression (2-148) of the self-energy in the SCBA. By defining the screened Coulomb interaction as

$$
\begin{aligned}
\approx\!\approx &= \sim\sim + \sim\!\bullet\!(P)\!\bullet\!\sim + \sim\!\bullet\!(P)\!\bullet\!\sim\!\bullet\!(P)\!\bullet\!\sim + \ldots \\
&= \sim\sim + \sim\!\bullet\!(P)\!\bullet\!\approx\!\approx,
\end{aligned}
\tag{2-175}
$$

with the term P, called the *polarizability*, given by the sum of all possible polarization diagrams

$$\text{(P)} = [\text{diagram}] + [\text{diagram}] + \ldots, \quad (2\text{-}176)$$

the diagrammatic expression of the carrier-carrier self-energy (2-148) can be written in the following form

$$\text{(}\Sigma\text{)} = [\text{diagram}] + [\text{diagram}] + [\text{diagram}] + \ldots. \quad (2\text{-}177)$$

Exchange

The second term in the right-hand side, which corresponds to a Fock diagram with the bare Coulomb interaction replaced by a screened interaction, contains the original Fock term, the direct collision term, as well as infinite higher-order terms containing various repetitions of polarization diagrams. In the *GW-approximation* [21, 47], the carrier-carrier self-energy is approximated by including the first two diagrams in (2-177), with the polarizability P computed in the lowest order approximation (first diagram in (2-176)). In terms of equations, the GW carrier-carrier self-energy reads [48] (using the simplified notation $\Sigma(\mathbf{r}_1, t_1; \mathbf{r}_2, t_2) = \Sigma(12)$)

$$\Sigma(12) = \Sigma_H(12) + \Sigma_{xc}(12), \quad (2\text{-}178)$$

where

$$\Sigma_H(12) = -\mathrm{i}\hbar\, \delta_C(t_1, t_2)\delta(\mathbf{r}_1 - \mathbf{r}_2) \int d\mathbf{r}_3\, V(\mathbf{r}_1 - \mathbf{r}_3) G(\mathbf{r}_3, t_1; \mathbf{r}_3, t_1) \quad (2\text{-}179)$$

is the Hartee self-energy, and

$$\Sigma_{xc}(12) = \mathrm{i}\hbar\, G(12)W(21), \quad (2\text{-}180)$$

is the exchange-correlation self-energy, with $W(12)$ the screened interaction, which satisfies the Dyson's equation (expressed diagrammatically in Eq. (2-175))

$$W(12) = V(12) + \int d3 \int d4\, V(13)P(34)W(42), \quad (2\text{-}181)$$

where we have introduced $V(12) = V(\mathbf{r}_1 - \mathbf{r}_2)\delta_C(t_1, t_2)$[21], and the polarizability function in the GW-approximation is given by

$$P(12) = -\mathrm{i}\hbar\, G(12)G(21). \quad (2\text{-}182)$$

[21] The delta function added to the bare coulomb potential takes into account that the interaction is instantaneous.

The Hartree term

Using the definition of the carrier density from Eq. (2-45), the Hartree term (2-179) can be written as

$$\begin{aligned}\Sigma_H(12) &= \delta_C(t_1,t_2)\delta(\mathbf{r}_1-\mathbf{r}_2)\int d\mathbf{r}_3\, V(\mathbf{r}_1-\mathbf{r}_3)n(\mathbf{r}_3,t_1)\\ &= \delta_C(t_1,t_2)\delta(\mathbf{r}_1-\mathbf{r}_2)\, U_{\text{Hartree}}(\mathbf{r}_1,t_1),\end{aligned} \tag{2-183}$$

where $U_{\text{Hartree}}(1)$ is the average potential produced by the carrier distribution $n(\mathbf{r},t)$. For the case of a Coulomb interaction

$$V(\mathbf{r}-\mathbf{r}') = \frac{e^2}{4\pi\epsilon|\mathbf{r}-\mathbf{r}'|}, \tag{2-184}$$

the Hartree potential can be demonstrated to be the solution of the Poisson equation

$$\frac{1}{e}\nabla\cdot\big(\epsilon(1)\nabla U_{Hartree}(1)\big) = \rho(1), \tag{2-185}$$

for a charge density $\rho(1) = -e\,n(1)$ (see Ref. [47], Sec. 2.8.1). When inserted in the differential form of the Dyson's equation (2-10a), the Dirac delta functions in (2-183) are cancelled by integration and the Hartree potential can be included in the left-hand side by defining the effective potential $U_{\text{eff}}(1) = U(1) + U_{\text{Hartree}}(1)$.

The exchange-correlation term

The exchange-correlation term of the carrier-carrier self-energy can be divided into the Fock and direct-collision self-energies by inserting Eq. (2-181) into (2-180)

$$\begin{aligned}\Sigma_{xc}(12) &= \Sigma_{\text{Fock}}(12) + \Sigma_{dc}(12)\\ &= \mathrm{i}\hbar\, G(12)V(21) + \mathrm{i}\hbar\, G(12)\int d3\int d4\, V(23)P(34)W(41).\end{aligned} \tag{2-186}$$

The Fock term, which has a singularity due to the presence of the Dirac delta function, introduces a correction potential due to exchange effects, similar to the case of the Hartree self-energy (its lesser and greater components are zero because of the delta function), while the direct collision term takes into account correlation effects between different particles. In the $\mathbf{k}\cdot\mathbf{p}$ basis (2-77), the covariant representation of the exchange-correlation self-energy can be obtained by inserting Eq. (2-180) into (2-68). Introducing the compact index notation $1=(i_1,m_1)$ for the basis functions $\{\phi_{1,\mathbf{k}}\}$, the discretized self-energy reads

$$\begin{aligned}\Sigma_{1,2}(\mathbf{k};t,t') &= \int d\mathbf{r}\int d\mathbf{r}'\,\phi^*_{1,\mathbf{k}}(\mathbf{r})\,\Sigma_{xc}(\mathbf{r},t;\mathbf{r}',t')\,\phi_{2,\mathbf{k}}(\mathbf{r}')\\ &= \mathrm{i}\hbar\int d\mathbf{r}\int d\mathbf{r}'\,\phi^*_{1,\mathbf{k}}(\mathbf{r})\,G(\mathbf{r},t;\mathbf{r}',t')W(\mathbf{r}',t';\mathbf{r},t)\,\phi_{2,\mathbf{k}}(\mathbf{r}'),\end{aligned} \tag{2-187}$$

where the GF can be expanded using (2-67), obtaining

$$\begin{aligned}\Sigma_{1,2}(\mathbf{k};t,t') &= \mathrm{i}\hbar \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{1,\mathbf{k}}(\mathbf{r}) \left(\sum_{\mathbf{k}'} \sum_{34} \phi_{3,\mathbf{k}'}(\mathbf{r}) G_{3,4}(\mathbf{k}';t,t') \phi^*_{4,\mathbf{k}'}(\mathbf{r}') \right) W(\mathbf{r}',t';\mathbf{r},t)\, \phi_{2,\mathbf{k}}(\mathbf{r}') \\ &= \mathrm{i}\hbar \sum_{\mathbf{k}'} \sum_{34} G_{3,4}(\mathbf{k}';t,t') \left(\int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{1,\mathbf{k}}(\mathbf{r}) \phi^*_{4,\mathbf{k}'}(\mathbf{r}') W(\mathbf{r}',t';\mathbf{r},t) \phi_{2,\mathbf{k}}(\mathbf{r}') \phi_{3,\mathbf{k}'}(\mathbf{r}) \right) \\ &= \mathrm{i}\hbar \sum_{\mathbf{k}'} \sum_{34} G_{3,4}(\mathbf{k}';t,t') W_{1,4,2,3}(\mathbf{k}'-\mathbf{k};t',t) \end{aligned} \tag{2-188}$$

with the matrix element of the screened interaction given by the four-index tensor

$$W_{1,2,3,4}(\mathbf{k}'-\mathbf{k};t,t') = \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{1,\mathbf{k}}(\mathbf{r}) \phi^*_{2,\mathbf{k}'}(\mathbf{r}')\, W(\mathbf{r},t;\mathbf{r}',t')\, \phi_{3,\mathbf{k}}(\mathbf{r}') \phi_{4,\mathbf{k}'}(\mathbf{r}). \tag{2-189}$$

Applying the Langreth rules, the real-time components of the GW self-energy (2-188) become

$$\Sigma^{\lessgtr}_{1,2}(\mathbf{k};t',t) = \mathrm{i}\hbar \sum_{\mathbf{q}} \sum_{34} G^{\lessgtr}_{3,4}(\mathbf{k}+\mathbf{q};t,t') W^{\gtrless}_{1,4,2,3}(\mathbf{q};t',t) \tag{2-190}$$

$$\begin{aligned}\Sigma^{R}_{1,2}(\mathbf{k};t,t') = \mathrm{i}\hbar \sum_{\mathbf{q}} \sum_{34} \big[& G^{<}_{3,4}(\mathbf{k}+\mathbf{q};t,t') W^{A}_{1,4,2,3}(\mathbf{q};t',t) \\ & + G^{R}_{3,4}(\mathbf{k}+\mathbf{q};t,t') W^{<}_{1,4,2,3}(\mathbf{q};t',t) \big], \end{aligned} \tag{2-191}$$

where the substitution $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ was made (if the transferred momentum $\mathbf{q}$ is small, the displacement of the summation region can be neglected). The steady-state expressions can be obtained by Fourier transforming with respect to the time difference $\tau = t - t'$, yielding

$$\Sigma^{\lessgtr}_{1,2}(\mathbf{k},E) = \mathrm{i} \sum_{\mathbf{q}} \sum_{34} \int \frac{dE'}{2\pi}\, G^{\lessgtr}_{3,4}(\mathbf{k}+\mathbf{q},E+E') W^{\gtrless}_{1,4,2,3}(\mathbf{q};E') \tag{2-192}$$

$$\begin{aligned}\Sigma^{R}_{1,2}(\mathbf{k},E) = \mathrm{i} \sum_{\mathbf{q}} \sum_{34} \int \frac{dE'}{2\pi}\, \big[& G^{<}_{3,4}(\mathbf{k}+\mathbf{q},E+E') W^{A}_{1,4,2,3}(\mathbf{q},E') \\ & + G^{R}_{3,4}(\mathbf{k}+\mathbf{q},E+E') W^{<}_{1,4,2,3}(\mathbf{q},E') \big]. \end{aligned} \tag{2-193}$$

Equation of motion of the screened interaction

The discretization of the polarization function (2-182) is straightforward as it just requires to expand the GFs into the basis $\{\phi_{1,\mathbf{k}}(\mathbf{r})\}$

$$\begin{aligned} P(12) &= -\mathrm{i}\hbar \left(\sum_{\mathbf{k}'} \sum_{57} \phi_{5,\mathbf{k}'}(\mathbf{r}_1) G_{5,7}(\mathbf{k}';t_1,t_2) \phi^*_{7,\mathbf{k}'}(\mathbf{r}_2) \right) \left(\sum_{\mathbf{k}} \sum_{68} \phi_{6,\mathbf{k}}(\mathbf{r}_2) G_{6,8}(\mathbf{k};t_2,t_1) \phi^*_{8,\mathbf{k}}(\mathbf{r}_1) \right) \\ &= \sum_{\mathbf{q}} \sum_{5678} \phi_{5,\mathbf{0}}(\mathbf{r}_1) \phi_{6,-\mathbf{q}}(\mathbf{r}_2) \underbrace{\left(-\mathrm{i}\hbar \sum_{\mathbf{k}'} G_{5,7}(\mathbf{k}';t_1,t_2) G_{6,8}(\mathbf{k}'-\mathbf{q};t_2,t_1) \right)}_{P_{5,6,7,8}(\mathbf{q};t_1,t_2)} \phi^*_{7,\mathbf{0}}(\mathbf{r}_2) \phi^*_{8,-\mathbf{q}}(\mathbf{r}_1), \end{aligned}$$

leading to the real-time, Fourier transformed expressions

$$P^{\lessgtr}_{5,6,7,8}(\mathbf{q},E') = -\mathrm{i}\sum_{\mathbf{k}'}\int \frac{dE''}{2\pi} G^{\lessgtr}_{5,7}(\mathbf{k}',E')G^{\gtrless}_{6,8}(\mathbf{k}'-\mathbf{q},E'-E'') \tag{2-194}$$

$$P^{R}_{5,6,7,8}(\mathbf{q},E') = -\mathrm{i}\sum_{\mathbf{k}'}\int \frac{dE''}{2\pi} [G^{R}_{5,7}(\mathbf{k}',E')G^{<}_{6,8}(\mathbf{k}'-\mathbf{q},E'-E'') + G^{<}_{5,7}(\mathbf{k}',E')G^{A}_{6,8}(\mathbf{k}'-\mathbf{q},E'-E'')]. \tag{2-195}$$

Likewise, the discretized Dyson's equation for the screened potential (2-181) is computed by substituting (2-181) into the tensor definition (2-189), and expanding the polarization function. The final real-time expressions, in steady-state conditions, yield [49]

$$W^{\lessgtr}_{1,2,3,4}(\mathbf{q},E') = \sum_{5678} W^{R}_{1,8,5,4}(\mathbf{q},E')P^{\lessgtr}_{5,6,7,8}(\mathbf{q},E')W^{A}_{7,2,3,6}(\mathbf{q},E') \tag{2-196}$$

$$W^{R}_{1,2,3,4}(\mathbf{q},E') = V_{1,2,3,4}(\mathbf{q}) + \sum_{5678} V_{1,8,5,4}(\mathbf{q})P^{R}_{5,6,7,8}(\mathbf{q},E')W^{R}_{7,2,3,6}(\mathbf{q},E'). \tag{2-197}$$

with the bare (Coulomb) interaction tensor

$$V_{1,2,3,4}(\mathbf{k}'-\mathbf{k}) = \int d\mathbf{r}\int d\mathbf{r}'\, \phi^{*}_{1,\mathbf{k}}(\mathbf{r})\phi^{*}_{2,\mathbf{k}'}(\mathbf{r}')\, V(\mathbf{r}-\mathbf{r}')\, \phi_{3,\mathbf{k}}(\mathbf{r}')\phi_{4,\mathbf{k}'}(\mathbf{r}). \tag{2-198}$$

The tensorial nature of the screened potential make its numerical implementation very tedious, in particular the computation of the retarded component (2-197), as the way in which is stated, it involves a double tensor product along with a self-consistent iterative cycle for W^R. Although it can be restated in the form (ignoring momentarily the indices)

$$W^{R}(\mathbf{q},E') = V(\mathbf{q})\epsilon_d(\mathbf{q},E')^{-1}, \tag{2-199}$$

with $\epsilon_d(\mathbf{q},E') = 1 - V(\mathbf{q})P^R(\mathbf{q},E')$ the dielectric function, it still involves a tensor inversion which becomes prohibitive for structures of few tens of nanometres long. In order to lower the computational cost, the usual solution is to compute W^R in the GW_0-approximation instead of solving Eq. (2-197) [48, 49], *i.e.*, the retarded component of the polarization function is approximated by the equilibrium polarization function $P^R \approx P^R_0 = G^R_0 G^<_0 + G^<_0 G^A_0$. Such an approximation is equivalent to assume that our electron system interacts with an homogeneous electron gas that remains at equilibrium. Although unphysical, this approximation have proven to give more accurate results than the full GW-approximation, since neglecting vertex corrections in the later (*i.e.*, higher-order diagrams in the polarization expansion) leads to inaccurate quasi-particle spectral properties [50, 51]. Under the assumption of small transferred momentum $\mathbf{q}$, the GW_0 retarded screened interaction, in the static limit, is given by

$$W^{R}_0(12) = \frac{e^2}{4\pi\epsilon|\mathbf{r}_1-\mathbf{r}_2|} e^{-q_0|\mathbf{r}_1-\mathbf{r}_2|}\,\delta(t_1-t_2), \tag{2-200}$$

with $q_0 = \sqrt{\frac{e^2 n}{\epsilon k_B T}}$ the inverse Debye-Hückel screening length for a screening medium with carrier density n. A derivation the previous formula can be found in [52], Sec. 3.5.6. Notice that, although W^R will be approximated according to (2-200), we will still compute $P^{\lessgtr}$ and $W^{\lessgtr}$ following the usual GW formulas (2-194) and (2-196).

The screened interaction tensor in the $\mathbf{k}\cdot\mathbf{p}$ basis

The screened interaction tensor, in the basis (2-77), is given by

$$\begin{aligned} W_{1,2,3,4}(\mathbf{k}'-\mathbf{k};t,t') &= \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{1,\mathbf{k}}(\mathbf{r})\phi^*_{2,\mathbf{k}'}(\mathbf{r}')\, W(\mathbf{r},t;\mathbf{r}',t')\, \phi_{3,\mathbf{k}}(\mathbf{r}')\phi_{4,\mathbf{k}'}(\mathbf{r}) \\ &= \frac{1}{A^2}\int d\mathbf{r}\int d\mathbf{r}'\, t_{i_1}(z)u^*_{m_1 0}(\mathbf{r})\, t_{i_2}(z')u^*_{m_2 0}(\mathbf{r}')e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{r}_t-\mathbf{r}'_t)} \\ &\times W(\mathbf{r},t;\mathbf{r}',t')\, t_{i_3}(z')u_{m_3 0}(\mathbf{r}')\, t_{i_4}(z)u_{m_4 0}(\mathbf{r}). \end{aligned} \tag{2-201}$$

Such as in the case of the bare Coulomb potential, the screened potential is assumed to be a long-range interaction that vary slowly over a lattice cell, so the Bloch function products can be approximated by lattice cell averages, and Eq. (2-201) becomes

$$\begin{aligned} W_{1,2,3,4}(\mathbf{k}'-\mathbf{k};t,t') &\approx \frac{1}{A^2}\int d\mathbf{r}\int d\mathbf{r}'\, t_{i_1}(z)t_{i_2}(z')e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{r}_t-\mathbf{r}'_t)}\, W(\mathbf{r},t;\mathbf{r}',t')\, t_{i_3}(z')t_{i_4}(z) \\ &\times \frac{1}{\Omega}\int_\Omega d\mathbf{r}\, u^*_{m_1 0}(\mathbf{r})u_{m_4 0}(\mathbf{r})\, \frac{1}{\Omega}\int_\Omega d\mathbf{r}'\, u^*_{m_2 0}(\mathbf{r}')u_{m_3 0}(\mathbf{r}') \\ &= \frac{1}{A^2}\int d\mathbf{r}\int d\mathbf{r}'\, t_{i_1}(z)t_{i_2}(z')e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot(\mathbf{r}_t-\mathbf{r}'_t)}\, W(\mathbf{r},t;\mathbf{r}',t')\, t_{i_3}(z')t_{i_4}(z) \\ &\times \delta_{m_1,m_4}\,\delta_{m_2,m_3}, \end{aligned} \tag{2-202}$$

where the change of variable $\bar{\mathbf{r}}_t = \mathbf{r}_t - \mathbf{r}'_t$ can be performed, and assuming that the screened interaction depends on the positions $\mathbf{r}$ and $\mathbf{r}'$ through their difference (this is true for W^R_0), expression (2-202) finally yields

$$\begin{aligned} W_{1,2,3,4}(\mathbf{k}'-\mathbf{k};t,t') &= \delta_{m_1,m_4}\,\delta_{m_2,m_3} \\ &\times \int dz \int dz'\, t_{i_1}(z)t_{i_2}(z')\, W(\mathbf{k}'-\mathbf{k};z-z';t,t')\, t_{i_3}(z')t_{i_4}(z), \end{aligned} \tag{2-203}$$

with

$$W(\mathbf{k}'-\mathbf{k};z-z';t,t') = \frac{1}{A}\int d\bar{\mathbf{r}}_t\, e^{\mathrm{i}(\mathbf{k}'-\mathbf{k})\cdot\bar{\mathbf{r}}_t}\, W(\mathbf{r},t;\mathbf{r}',t'), \tag{2-204}$$

the 2D Fourier transform of the screened potential with respect to the transversal position vector difference $\bar{\mathbf{r}}_t = \mathbf{r}_t - \mathbf{r}'_t$. For the case of the statically screened retarded interaction

(2-200), after Fourier transforming with respect to the time coordinate, Eq. (2-204) gives

$$W_0^R(\mathbf{k}' - \mathbf{k}; z - z'; E') = \frac{e^2}{2\epsilon A} \frac{e^{-\sqrt{(\mathbf{k}'-\mathbf{k})^2+q_0^2}\,|z-z'|}}{\sqrt{(\mathbf{k}'-\mathbf{k})^2 + q_0^2}}, \tag{2-205}$$

and the matrix element (2-203) yields

$$\begin{aligned} W_{1,2,3,4}^R(\mathbf{k}' - \mathbf{k}, E') = & \frac{e^2}{2\epsilon A\sqrt{(\mathbf{k}'-\mathbf{k})^2 + q_0^2}}\, \delta_{m_1,m_4}\, \delta_{m_2,m_3} \\ & \times \int dz \int dz'\, t_{i_1}(z) t_{i_2}(z')\, e^{-\sqrt{(\mathbf{k}'-\mathbf{k})^2+q_0^2}\,|z-z'|}\, t_{i_3}(z') t_{i_4}(z), \end{aligned} \tag{2-206}$$

which can be treated as a pre-computable quantity.

2.7.4 Büttiker probes

The Büttiker probes were originally proposed as a phenomenological model to provide a simplified description of the relevant (intraband) self-energies, *e.g.*, electron-electron and electron-phonon scattering, which are usually obtained within the self-consistent Born approximation through a fixed-point iterative solution of the NEGF equations [53, 54, 55]. Within the Büttiker formalism, virtual contacts or *probes* are placed at a given position $\mathbf{r}_p$, resulting in the local self-energy [56, 57]

$$\Sigma_{BP}^R(\mathbf{r}, \mathbf{r}, E) = -\frac{\mathrm{i}\hbar}{2\tau_s} \delta(\mathbf{r} - \mathbf{r}_p), \tag{2-207}$$

with $\tau_s \equiv \tau_s(E)$ the (energy-dependent) scattering time for the p-th probe, related to some scattering mechanism s. The lesser and greater components of the self-energy are obtained by assuming the quasi-equilibrium relations (2-42) and (2-43), given by the fluctuation-dissipation theorem

$$\Sigma_{BP}^<(\mathbf{r}, \mathbf{r}, E) = \mathrm{i} f_{FD}(E - E_{BP}) \Gamma_{BP}(\mathbf{r}, \mathbf{r}, E) \tag{2-208}$$

$$\Sigma_{BP}^>(\mathbf{r}, \mathbf{r}, E) = \mathrm{i}\big[f_{FD}(E - E_{BP}) - 1\big] \Gamma_{BP}(\mathbf{r}, \mathbf{r}, E), \tag{2-209}$$

where $\Gamma_{BP} = \mathrm{i}\big[\Sigma_{BP}^R - \Sigma_{BP}^A\big]$, and we have defined the Fermi-level $E_{BP} \equiv E_{BP}(\mathbf{r})$ of the p-th probe. Current conservation is achieved by simply adjusting the position-dependent Fermi levels of the probes. The numerical and conceptual simplicity of the Büttiker probes does not limit their application to phenomenological and oversimplified dissipation models, as was shown in the context of quantum cascade lasers [58]. Recently, carrier-phonon self-energies compatible with the Büttiker-probe formalism were employed in atomistic simulation of graphene nanotransistors [56]. For the case of acoustic phonons, an expression for the retarded self-energy consistent with the Büttiker-probe formalism, can be obtained from

the acoustic deformation potential self-energy (2-155). The quantity $1/\Delta_{ij}$ in Eq. (2-156) is a very narrow function peaked at $x_i = x_j$, that can be well approximated by $1/\Delta_{ij} \approx \delta_{i,j} 1/a_L$ [31]. With this approximation, the local acoustic deformation potential self-energy reads

$$\mathbf{\Sigma}^R_{\mathrm{aco},ii}(\mathbf{k}, E) = \frac{D_a^2 k_B T}{A\rho u_l^2 a_L} \sum_{\mathbf{k}'} \left\{ \mathbf{S}\mathbf{G}^R(\mathbf{k}', E)\mathbf{S} \right\}_{i,i}, \tag{2-210}$$

where we have substituted $\mathbf{k}' = \mathbf{k} - \mathbf{q}_t$. An optical deformation potential phonon scattering self-energy can be obtained by inserting the optical deformation potential model (see [59], Sec. 3.4)

$$|U_{\mathbf{q}}| = \sqrt{\frac{\hbar D_{tK}^2}{2\rho\,\omega_0}}, \tag{2-211}$$

in Eqs. (2-152) and (2-153). Unlike the Fröhlich model for polar optical scattering, this scattering strength term is independent of the phonon wavevector, which simplifies the integration over the reciprocal space. Following Ref. [56], deformation potential scattering assisted by optical phonons results in a (local) intraband self-energy whose retarded component is given by

$$\begin{aligned} \mathbf{\Sigma}^R_{\mathrm{op},ii}(\mathbf{k}, E) = \frac{\hbar D_{tK}^2}{2A\rho\omega_0 a_L} \sum_{\mathbf{k}'} & \left[n_0 - f_{\mathrm{FD}}\left(E - \hbar\omega_0 - E_{\mathrm{BP}}^{(m)}\right) + 1 \right] \left\{ \mathbf{S}\mathbf{G}^R(\mathbf{k}', E - \hbar\omega_0)\mathbf{S} \right\}_{i,i} \\ & + \left[n_0 + f_{\mathrm{FD}}\left(E + \hbar\omega_0 - E_{\mathrm{BP}}^{(m)}\right) \right] \left\{ \mathbf{S}\mathbf{G}^R(\mathbf{k}', E + \hbar\omega_0)\mathbf{S} \right\}_{i,i}, \end{aligned} \tag{2-212}$$

where f_{FD} is the Fermi-Dirac function, $E_{\mathrm{BP}}^{(m)}(z_i)$ is the position dependent Fermi level of the Büttiker probes in the band m, $D_{t,K}$ is the optical deformation potential, ρ is the mass density of the material, a_L is the lattice constant, and n_0 is the phonon occupation number for a population of dispersionless optical phonons with effective energy $\hbar\omega_0$.

2.8 Poisson equation

The last ingredient in a NEGF framework for optoelectronic device simulation is the Poisson equation, whose solution gives the carrier-carrier interaction term at the Hartree level, as was already pointed out in Sec. 2.7.3. The Hartree potential $U_{\mathrm{Hartree}}(z) = -e\varphi(z)$ entering the Dyson's equation is given by the Poisson equation

$$\frac{\partial}{\partial z}\left(\epsilon(z) \frac{\partial \varphi}{\partial z}(z) \right) = -\rho(z), \tag{2-213}$$

where φ is the electrostatic potential, ϵ is the static dielectric constant, and the charge density distribution is given by

$$\rho(z) = e\big[p(z) - n(z) + N_D^+(z) - N_A^-(z)\big] + \rho_{\mathrm{fix}}(z). \tag{2-214}$$

Here n and p are the electron and hole densities, and N_D^+ and N_A^- are the density of ionized donor and acceptor dopants, respectively. Additional fixed charge densities, *e.g.*, arising from piezoelectric or spontaneous polarization effects, are taken into account in ρ_{fix}. The carrier densities have a nonlinear dependency on φ through the relations

$$n(z) = N_C \, \mathcal{F}_{1/2} \left(\frac{E_F^{(n)}(z) - E_{C,0}(z) + e\varphi(z)}{k_B T} \right) \tag{2-215a}$$

$$p(z) = N_V \, \mathcal{F}_{1/2} \left(\frac{E_{V,0}(z) - e\varphi(z) - E_F^{(p)}(z)}{k_B T} \right), \tag{2-215b}$$

where N_C and N_V are the three-dimensional effective electron and hole densities of states, $E_{C,0}$ and $E_{V,0}$ are the conduction and valence band edges at flat-band conditions, and $E_F^{(n)}$ and $E_F^{(p)}$ are the electron and hole quasi-Fermi levels, which are computed from the NEGF electron and hole densities (2-82) and (2-83), respectively, via the formulas

$$E_F^{(n)}(z) = E_{C,0}(z) - e\varphi(z) + k_B T \, \mathcal{F}_{-1/2} \left(\frac{n_{\text{NEGF}}(z)}{N_C} \right) \tag{2-216a}$$

$$E_F^{(p)}(z) = E_{V,0}(z) - e\varphi(z) - k_B T \, \mathcal{F}_{-1/2} \left(\frac{p_{\text{NEGF}}(z)}{N_V} \right), \tag{2-216b}$$

with $\mathcal{F}_\eta$ the Fermi-Dirac integral of order η [60, 61]. The Poisson equation is discretized using a finite-element implementation[22] and is then solved by Newton iteration. Open boundaries are imposed by requiring the electrostatic potential to be equal to the potential of the equilibrium contacts, *i.e.*, Eq. (2-213) is solved using the Dirichlet boundary conditions

$$\varphi(z_1) = \varphi_L, \qquad \varphi(z_N) = \varphi_R, \tag{2-217}$$

where φ_L and φ_R are the electrostatic potentials of the left and right contacts, respectively. In order to find the contact potentials, first we need to solve the Poisson equation at thermodynamic equilibrium, by imposing vanishing electric field and zero net charge at the contacts, *i.e.*, $\frac{\partial \varphi}{\partial z}(z) = 0$ and $\rho(z) = 0$ for $z = z_1$, $z = z_N$. The equilibrium boundary potentials $\varphi_{L,0}$ and $\varphi_{R,0}$ are then used to determine φ_L and φ_R for an externally applied bias voltage V_{bias} through the relations

$$\varphi_L = \varphi_{L,0} + V_{\text{bias}}, \qquad \varphi_R = \varphi_{R,0}. \tag{2-218}$$

The band edge energies and contact Fermi levels, necessary for the computation of the boundary self-energies, are obtained from the solution of the Poisson equation as follows

$$E_C(z) = E_{C,0}(z) - e\varphi(z) \tag{2-219}$$

$$E_V(z) = E_{V,0}(z) - e\varphi(z), \tag{2-220}$$

[22] A detailed description of the numerical implementation using a finite-element discretization scheme can be found in [30], Sec. 2.4.4.

and

$$\mu_L = E_{C,0}(z_1) - e\varphi_L + k_B T\,\mathcal{F}_{-1/2}\left(\frac{n(z_1)}{N_C}\right) \tag{2-221a}$$

$$\mu_R = E_{C,0}(z_N) - e\varphi_R + k_B T\,\mathcal{F}_{-1/2}\left(\frac{n(z_N)}{N_C}\right). \tag{2-221b}$$

Eq. (2-213) together with the carrier densities from Eqs. (2-82) and (2-83), form a closed set of equations that can be solved self-consistently in what is usually called the *outer-loop*. The name comes from the fact that each iteration of the outer-loop involves the solution of another self-consistent set of equations, *i.e.*, the Dyson's and Keldysh equations (2-37a) and (2-37b), respectively, along with the scattering self-energies in the SCBA, which is usually referred as the *inner-loop*. More details about the numerical implementation of the inner and the outer loops will be discussed in the next section.

2.8.1 Ionized dopant concentrations

For a single dopant species, the concentration of ionized shallow donors and acceptors is given by the expressions [33]

$$N_A^-(z) = N_A\left[1 + g_A \exp\left(\frac{E_V(z) + \Delta E_A - E_F^{(p)}(z)}{k_B T}\right)\right]^{-1} \tag{2-222a}$$

$$N_D^+(z) = N_D\left[1 + g_D \exp\left(\frac{E_F^{(n)}(z) - E_C(z) + \Delta E_D}{k_B T}\right)\right]^{-1}, \tag{2-222b}$$

where N_A and N_D are the acceptor and donor dopant concentrations, ΔE_A and ΔE_D are the acceptor and donor ionization energies, and $g_A = 4$ and $g_D = 2$ are the ground state degeneracy factors for acceptor and donor levels, respectively.

2.8.2 Polarization charges

Variations in the polarization field in polar materials, or as a result of piezoelectric effect in strained materials, can produce strong electric fields that may significantly impact the transport properties of a device. Such variations can occur in nanostructures, *e.g.*, because of the presence of an heterostructure, a quantum-well or because of a mismatch of the lattice parameter with respect to the substrate. In this work we will only consider the polarization fields for the case of III-N semiconductors. According to the model proposed by *Bernardini* and *Fiorentini* [62], the spontaneous polarization (in units C m^{-2}) along the growth z-axis is calculated as

$$P_{sp}(\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}) = -0.090x - 0.034(1-x) + 0.019x(1-x) \tag{2-223}$$

$$P_{sp}(\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}) = -0.042x - 0.034(1-x) + 0.038x(1-x). \tag{2-224}$$

Similarly, piezoelectric polarization can be expressed as (see Ref. [33], App. D)

$$P_{pz}(\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{N}) = -0.0525x + 0.0282x(1-x) \tag{2-225}$$
$$P_{pz}(\mathrm{In}_x\mathrm{Ga}_{1-x}\mathrm{N}) = 0.148x - 0.0424x(1-x). \tag{2-226}$$

In all the previous formulas, x represents the alloy concentration. The total polarization $P = P_{sp} + P_{pz}$ then enters the Poisson equation through the polarization charges

$$\rho_{\mathrm{pol}}(z) = -\frac{\partial P}{\partial z}(z). \tag{2-227}$$

A more detailed discussion on the subject can be found in Sec. 11.5.3 of [33], and its numerical implementation in a finite-element framework in Sec. 2.4.3 of [30].

2.9 Numerical considerations and choice of simulation parameters

The NEGF simulation flow, showing iteration steps for both inner and outer-loop, is described in Fig. **2-4**. The inner self-consistency loop connects the equations for the GFs and self-energies, while the outer loop provides the update of the Hartree potential from the solution of Poisson's equation. An initial condition is used at the first step of the process, which is chosen to be the noninteracting GF $\mathbf{G}_0$, *i.e.*, at ballistic conditions, including only the boundary self-energy. A *drift-diffusion* simulation is used to determine the initial values of the electrostatic potential and band edge energies.

Despite being a very powerful method for the simulation of nanoscale optoelectronic devices, the NEGF approach is very computationally demanding. If we for example consider a small resistor of 50 nm length, discretized using a uniform mesh spacing of $\Delta z = 0.5$ nm, this would correspond to a total number of $N = 101$ discretization points. Additionally, if a discretization in energy and k-space is chosen such that $N_E = 300$ and $N_k = 50$ points are considered, respectively, using a 4-band $\mathbf{k}\cdot\mathbf{p}$ model so $N_b = 4$, then each single GF or self-energy would require to store $(N_b \times N)^2 \times N_E \times N_k = 2.448 \times 10^9$ complex numbers[23], which amounts to approximately 39 GB per operator using a double-precision floating point format. This simple example is useful to see how intensive the NEGF method is, both in therms of time, because of all the matrix multiplications and inversions it involves and which should be iterated to convergence, and in terms of memory, since at least $G^<$, G^R, $\Sigma^<$ and Σ^R should be stored at each iteration. The limitations illustrated above impose a restriction on the total length of the structures that can be analysed using the NEGF approach, usually

[23] Exploiting the symmetry properties in Eqs. (2-39a)-(2-39b), the total space needed for the storage of the operators can be reduced by half, since only the upper or lower triangular parts of each matrix are necessary to recover the full operator.

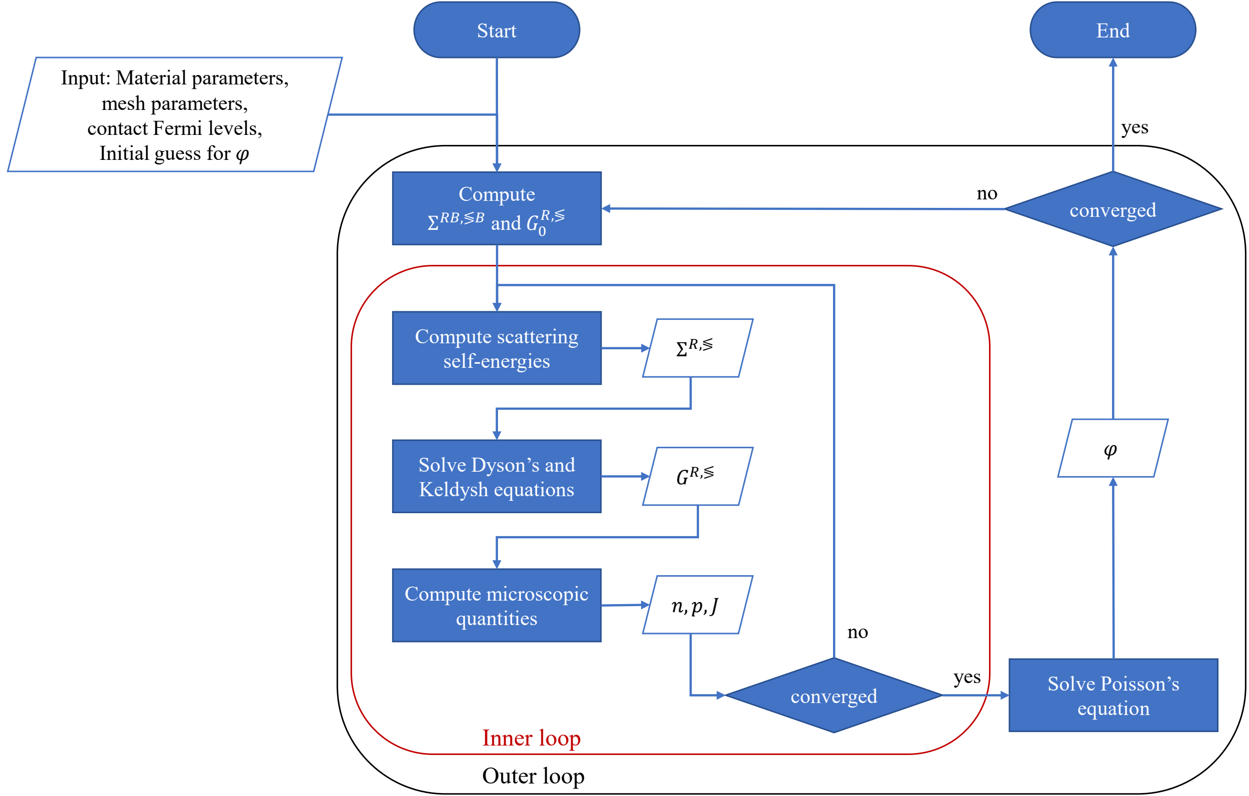


Figure 2-4: *Flowchart for the calculation of physical quantities from the Green's functions. The inner loop connects the Green's functions and scattering self-energies in the SCBA. Convergence is achieved when the relative change of the total current falls below some residual value ϵ_{inner}, see Eq.* (2-229)*. The outer loop provides the update of the electrostatic potential φ from the solution of the Poisson's equation.*

applied to devices with no more than 80 nm length. Nevertheless, some improvements can be implemented to the NEGF method so it can be applied to longer structures, *e.g.*, in terms of convergence with the *Pulay method*, or in terms of memory requirements using *mode order reduction*, which were already discussed in a previous work [63].

One key aspect before running the NEGF simulations is choosing appropriately all the simulation parameters, such as ΔE, Δk, $k_{\max}$, and the energy grid boundaries $E_{\min}$ and $E_{\max}$. Energy and transverse momentum resolutions ΔE and Δk, respectively, should be chosen small enough to be able to recognize all the spectral features of the simulated device, specially if the structure contains, *e.g.*, a quantum-well introducing very sharp features in the density of states, since spectral quantities are going to be integrated in energy and momentum. Nonetheless, a good compromise between the resolution and the resulting number of discretization points should be found, so they allow a reasonable computation time. In the case of transverse momentum, effects of choosing high values of Δk are evidenced as

oscillating peaks in the spectrally resolved quantities. Here we take the value suggested in [23] for this quantity $\Delta k \approx 0.003\pi/a_L$, with a_L the lattice parameter. In the case of the energy resolution, a value of ΔE less than 10 meV seems to give a good compromise between resolution and computation time.

On the other hand, the spectral range (both in energy and transversal momentum) should be chosen such that all the important phenomena is included in the simulation. Good values for E_{min} and E_{max} are about 0.5-1.0 eV below and above the band edge extrema. In most cases, energies in the bandgap region, between the conduction band minimum and the valence band maximum (minus/plus few tens of meV) may also be neglected, reducing further the number of energy discretization points. Finally, the value of k_{max} should be chosen according to parameters like the scattering strength of the different mechanisms included in the simulation, the contact doping, and the effective masses of the materials that compose the device. In III-N semiconductors, featuring very high heavy-hole masses, the maximum transverse momentum should be chosen particularly large. Some values of k_{max} have been suggested in the literature [31, 23], from which $k_{max} \approx 0.14\pi/a_L$ seems to be a more than reasonable value in most cases, hence obtaining an average of $N_k = k_{max}/\Delta k \approx 50$ points for the transverse momentum discretization.

Among other parameters that need to be adjusted, the polar optical phonon energy $\hbar\omega_{\rm LO}$ will be assumed to be constant regardless on the material, equal to 90 meV for III-N compounds, and 30 meV for GaSb/AlSb/InAs and its compounds. This assumption apart of simplifying the implementation of the code, will help to avoid GF interpolation when computing the polar optical phonon self-energy, by taking ΔE such that $\hbar\omega_{\rm LO}/\Delta E = k$ integer, so energy translations of $\hbar\omega_{\rm LO}$ on any point of the grid will end up in another point on the grid.

As a last remark, it is important to clarify how the convergence of the inner and outer loop will be treated. For the outer-loop, the Poisson equation will be solved until the potential change at the iteration k falls below a residual value $\epsilon_{\rm outer}$, *i.e.*,

$$\delta\varphi_k = |\varphi_k - \varphi_{k-1}| < \epsilon_{\rm outer}. \tag{2-228}$$

Regarding the inner iterations, these will be performed until the relative change in the device total current at the iteration k goes below a limiting value $\epsilon_{\rm inner}$, *i.e.*,

$$\frac{|\mathbf{J}_k - \mathbf{J}_{k-1}|}{|\mathbf{J}_k|} < \epsilon_{\rm inner}. \tag{2-229}$$

In all the following simulations $\epsilon_{\rm inner} = 10^{-6}$ and $\epsilon_{\rm outer} = 10^{-3}$ will be assumed.

Chapter 3

Modeling of trap-assisted tunneling in GaN based LEDs

This chapter will be devoted to the analysis of defect-mediated recombination in the transport properties of GaN-based LEDs in the forward-bias regime. Experimental studies have shown evidence relating the presence of defects to the efficiency droop effect, i.e. the decline of internal quantum efficiency at high current levels [64, 8]. Instead of focusing on the optical emission spectrum, our goal will be to study the subthreshold *I-V* characteristics of GaN-based LEDs, in which high ideality factors have been attributed to trap-assisted tunneling (TAT), and hence they constitute a signature of the presence of defects in the material [12, 13, 15]. In order to achieve this, the first step will consist in developing a quantum kinetic treatment of trap-assisted tunneling in the framework of the nonequilibrium Green's function approach, expressed in our $\mathbf{k}\cdot\mathbf{p}$ formulation. Few attempts have been performed before to develop a quantum kinetic picture of Shockley-Read-Hall recombination with NEGF, from which we highlight the recent work of *U. Aeberhard* in [65], which derives a self-energy for non-radiative transitions based on the multiphonon theory proposed by *A. Schenk* in [66]. This multiphonon SRH self-energy will be our starting point, and will be used to confront numerically the *I-V* characteristics of few devices, from which some insight will be extracted on the role of defect-mediated tunneling and recombination.

3.1 NEGF treatment of localized deep defects

Strongly localized point defects in semiconductors generate confined energy levels inside the gap that favor carrier recombination by serving as a ladder for subsequent thermal relaxation, which is often enhanced by the presence of an electric field leading to tunneling processes. Electronic states of a point defect, characterized by a single (ground) energy level ε_d, can be

described by the field operator [65]

$$\hat{\Psi}(\mathbf{r},t) = \psi(\mathbf{r})\hat{d}(t), \tag{3-1}$$

where $\hat{d}$ is the annihilation operator for an electron at the energy level ε_d, and $\psi(\mathbf{r})$ is the defect wave function, which is assumed to describe the probability amplitude of an electron localized in a volume Ω_d around a position $\mathbf{r}_d$, and is therefore approximated by

$$\psi(\mathbf{r}) \approx c_d\sqrt{\Omega_d}\,\delta(\mathbf{r}-\mathbf{r}_d), \tag{3-2}$$

with $c_d = 1$ for nondegenerate states, and the factor $\sqrt{\Omega_d}$ is added for normalization[1]. The Green's function for the single point defect system is then given by

$$\begin{aligned} G_d(\mathbf{r},t;\mathbf{r}',t') &= -\frac{\mathrm{i}}{\hbar}\langle \hat{T}_C\{\hat{\Psi}(\mathbf{r},t)\hat{\Psi}^\dagger(\mathbf{r}',t')\}\rangle \\ &= -\frac{\mathrm{i}}{\hbar}\langle \hat{T}_C\{\hat{d}(t)\hat{d}^\dagger(t')\}\rangle\,\Omega_d\,\delta(\mathbf{r}-\mathbf{r}_d)\,\delta(\mathbf{r}'-\mathbf{r}_d) \\ &= \Omega_d\,\delta(\mathbf{r}-\mathbf{r}_d)\,\delta(\mathbf{r}'-\mathbf{r}_d)\,G_d(t,t'). \end{aligned} \tag{3-3}$$

In steady-state conditions, the real-time components of the single level defect Green's function $G_d(t,t')$ satisfy the Dyson's and Keldysh equations, which can be written in energy coordinates

$$G_d^R(E) = [E - \varepsilon_d - \Sigma_d^R(E)]^{-1} \tag{3-4a}$$

$$G_d^{\lessgtr}(E) = G_d^R(E)\Sigma_d^{\lessgtr}(E)G_d^A(E), \tag{3-4b}$$

with the defect self-energy Σ_d encoding the coupling of the confined defect state with the extended states in both the conduction and valence bands, resulting in a broadening of the confined level due to the escape of carriers. Following [66], we assume that the capture (emission) process of electrons into (from) the confined defect state is enabled thanks to electron-phonon interactions, and therefore the defect self-energy should be given by an expression similar to (2-153). As usual, the starting point in order to find an expression for Σ_d is the perturbation expansion of the GF that describe the evolution of the device-defect system, whose second-order contribution is given by (B-5)

$$G^{(2)}(12) = \left(\frac{-i}{\hbar}\right)^3 \frac{1}{2}\,\langle \hat{T}_C\left\{\int d3\int d4\hat{\Psi}^\dagger(3)V_{ep}(3)\hat{\Psi}(3)\hat{\Psi}^\dagger(4)V_{ep}(4)\hat{\Psi}(4)\hat{\Psi}(1)\hat{\Psi}^\dagger(2)\right\}\rangle, \tag{3-5}$$

where we have substituted the interaction Hamiltonian (B-3), and V_{ep} is the electron-phonon interaction potential (B-1). At this point, Wick's theorem is applied to factor the field

[1] In order for the probability distribution $|\psi(\mathbf{r})|^2$ to make sense, the delta Dirac is intended here as the square function $\bar{\delta}(\mathbf{r}-\mathbf{r}_d) = 1/\Omega_d$ for $\mathbf{r}\in\Omega_d$, and zero otherwise. This ensures that $\int d\mathbf{r}\,|\psi(\mathbf{r})|^2 = 1$, and since Ω_d is much smaller than the mesh spacing in our basis, the function $\bar{\delta}$ can be approximated as the usual Dirac delta when integrated with other functions.

operator product. Restricting the resulting terms to those corresponding to Fock diagrams (there are two equivalent factorizations), yields

$$\begin{aligned} G^{(2)}(12) &= \left(\frac{-i}{\hbar}\right)^3 \int d3 \int d4\, \langle \hat{T}_C \{V_{ep}(3) V_{ep}(4)\}\rangle \\ &\times \langle \hat{T}_C \left\{ \hat{\Psi}(1)\hat{\Psi}^\dagger(3) \right\}\rangle \langle \hat{T}_C \left\{ \hat{\Psi}(3)\hat{\Psi}^\dagger(4) \right\}\rangle \langle \hat{T}_C \left\{ \hat{\Psi}(4)\hat{\Psi}^\dagger(2) \right\}\rangle \\ &= \int d3 \int d4\, G_0(13)\, \langle \hat{T}_C \{V_{ep}(3) V_{ep}(4)\}\rangle\, G_0(34)\, G_0(42). \end{aligned} \tag{3-6}$$

Since we are interested in a particle at the defect state, we assume $\mathbf{r}_1, \mathbf{r}_2 = \mathbf{r}_d$, so the noninteracting propagators $G_0(13)$ and $G_0(42)$ can be approximated by the defect GF (3-3)[2], leading to the expression

$$\begin{aligned} G^{(2)}(12) &= \Omega_d\, \delta(\mathbf{r}_1 - \mathbf{r}_d)\, \delta(\mathbf{r}_2 - \mathbf{r}_d) \\ &\times \int_C dt_3 \int_C dt_4\, G_0(t_1, t_3)\, V \langle \hat{T}_C \{V_{ep}(\mathbf{r}_d, t_3) V_{ep}(\mathbf{r}_d, t_4)\}\rangle\, G_0(\mathbf{r}_d, t_3; \mathbf{r}_d, t_4) G_0(t_4, t_2) \\ &= \Omega_d\, \delta(\mathbf{r}_1 - \mathbf{r}_d)\, \delta(\mathbf{r}_2 - \mathbf{r}_d) \int_C dt_3 \int_C dt_4\, G_0(t_1, t_3)\, \Sigma_d(t_3, t_4) G_0(t_4, t_2), \end{aligned} \tag{3-7}$$

from which the defect self-energy, in the SCBA, can be identified as

$$\begin{aligned} \Sigma_d(t, t') &= \Omega_d\, \langle \hat{T}_C \{V_{ep}(\mathbf{r}_d, t) V_{ep}(\mathbf{r}_d, t')\}\rangle\, G(\mathbf{r}_d, t; \mathbf{r}_d, t') \\ &= \Omega_d\, \langle \hat{T}_C \{V_{ep}(\mathbf{r}_d, t) V_{ep}(\mathbf{r}_d, t')\}\rangle \\ &\times \frac{1}{A} \sum_{\mathbf{k}} \sum_{ijmn} t_i(z_d)\, u_{m,0}(\mathbf{r}_d)\, G_{im,jn}(\mathbf{k}; t, t')\, t_j(z_d)\, u^*_{n,0}(\mathbf{r}_d), \end{aligned} \tag{3-8}$$

where we have expanded the GF in the basis (2-77), and z_d is the longitudinal component of the defect position vector $\mathbf{r}_d$. For simplicity z_d is assumed to lie on the spatial discretization mesh, allowing to exploit the identity $t_i(z_d) = \delta_{i,i_d}$, where i_d is the mesh point corresponding to z_d, and the Bloch function product is approximated as a lattice cell average around $\mathbf{r}_d$, obtaining the final expression

$$\Sigma_d(t, t') = \Omega_d\, \langle \hat{T}_C \{V_{ep}(\mathbf{r}_d, t) V_{ep}(\mathbf{r}_d, t')\}\rangle\, \frac{1}{A} \sum_{\mathbf{k}} \mathrm{Tr}[\mathbf{G}_{i_d, i_d}(\mathbf{k}; t, t')], \tag{3-9}$$

where the trace is over band indices. Using the definitions of the electron-phonon potential

[2] In the case when $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}_d$, $G_0(42)$ describe the free propagation of a particle from the defect at $\mathbf{r}_d$ to an arbitrary position $\mathbf{r}_4$. Since no scattering is considered, the particle should remain confined at the defect position, *i.e.*, $G_0(42) = 0$ for any $\mathbf{r}_4 \neq \mathbf{r}_d$, thus justifying the approximation. The same reasoning can be applied to $G_0(13)$.

(B-1) and the free phonon propagator (B-10), the ensemble average in Eq. (3-9) gives

$$\begin{aligned}\langle \hat{T}_C \{V_{ep}(\mathbf{r}_d,t)V_{ep}(\mathbf{r}_d,t')\}\rangle &= \mathrm{i}\hbar \sum_{l,l'\geq 0}\sum_{\mathbf{q},\mathbf{q}'} M(l,\mathbf{q})M(l',\mathbf{q}')e^{\mathrm{i}(\mathbf{q}+\mathbf{q}')\cdot\mathbf{r}_d} D_0(l,\mathbf{q};t,t')\delta_{\mathbf{q},-\mathbf{q}'}\,\delta_{l,l'} \\ &= \mathrm{i}\hbar \sum_{l\geq 0}\sum_{\mathbf{q}} M(l,\mathbf{q})M(l,-\mathbf{q})D_0(l,\mathbf{q};t,t') \\ &= \mathrm{i}\hbar \sum_{l\geq 0}\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 D_0(l,\mathbf{q};t,t'). \end{aligned} \tag{3-10}$$

Here l denotes the number of phonon modes participating in the multiphonon transition. The lesser and greater components of the defect self-energy (3-9) can be obtained using the Langreth rules (see [20], Sec. 4.3), obtaining the real-time expressions

$$\Sigma_d^{\lessgtr}(t,t') = \mathrm{i}\hbar\,\Omega_d \sum_{l\geq 0}\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 D_0^{\lessgtr}(l,\mathbf{q};t,t')\,\frac{1}{A}\sum_{\mathbf{k}} \mathrm{Tr}[\mathbf{G}^{\lessgtr}_{i_d,i_d}(\mathbf{k};t,t')], \tag{3-11}$$

which can be Fourier transformed with respect to the time difference $\tau = t - t'$, leading to the steady-state expressions

$$\Sigma_d^{\lessgtr}(E) = \mathrm{i}\Omega_d \sum_{l\geq 0}\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 \frac{1}{A}\sum_{\mathbf{k}} \int \frac{dE'}{2\pi} D_0^{\lessgtr}(l,\mathbf{q},E')\,\mathrm{Tr}[\mathbf{G}^{\lessgtr}_{i_d,i_d}(\mathbf{k},E-E')]. \tag{3-12}$$

Assuming an equilibrium phonon population, characterized by a dispersion-less phonon energy $\hbar\omega_0$, Eq. (3-12) finally yields

$$\Sigma_d^{<}(E) = +2\pi\mathrm{i}\sum_{l\geq 0}\left[\mathcal{M}_d^{\mathrm{capt}}(l)\,n(z_d,E+\hbar\omega_l) + \mathcal{M}_d^{\mathrm{em}}(l)\,n(z_d,E-\hbar\omega_l)\right] \tag{3-13a}$$

$$\Sigma_d^{>}(E) = -2\pi\mathrm{i}\sum_{l\geq 0}\left[\mathcal{M}_d^{\mathrm{capt}}(l)\,p(z_d,E-\hbar\omega_l) + \mathcal{M}_d^{\mathrm{em}}(l)\,p(z_d,E+\hbar\omega_l)\right], \tag{3-13b}$$

where $\hbar\omega_l = l\hbar\omega_0$ is the energy of the multiphonon transition, and we have used the definition of the lesser and greater free phonon GF $D_0^{\lessgtr}$ in (B-33a). This work will assume the effective phonon energy $\hbar\omega_0$ to be equal to the polar optical phonon energy $\hbar\omega_{\mathrm{LO}}$. In Eqs. (3-13a) and (3-13b) we have defined the (squared) matrix elements (with units $\mathrm{J}^2\,\mathrm{m}^3$)

$$\mathcal{M}_d^{\mathrm{capt}}(l) = \Omega_d\,(n_l+1)\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 \tag{3-14a}$$

$$\mathcal{M}_d^{\mathrm{em}}(l) = \Omega_d\,n_l\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2, \tag{3-14b}$$

associated with the multiphonon capture (emission) process of carriers from (into) the extended states into (from) the defect state, which satisfy the relation

$$\frac{\mathcal{M}_d^{\mathrm{em}}(l)}{\mathcal{M}_d^{\mathrm{capt}}(l)} = \frac{n_l}{n_l+1} = e^{-\beta\hbar\omega_l}, \tag{3-15}$$

with

$$n_l = \left(e^{\beta\hbar\omega_l} - 1\right)^{-1},$$

the occupation number of a phonon in mode l, and $\beta^{-1} = k_B T$. In Eqs. (3-13a)-(3-13b), the definition of the spectrally resolved carrier densities from Eqs. (2-48a)-(2-48b) was used

$$n(z_d, E) = \frac{-\mathrm{i}}{2\pi A}\sum_{\mathbf{k}} \mathrm{Tr}[\mathbf{G}^{<}_{i_d,i_d}(\mathbf{k}, E)] \tag{3-16a}$$

$$p(z_d, E) = \frac{\mathrm{i}}{2\pi A}\sum_{\mathbf{k}} \mathrm{Tr}[\mathbf{G}^{>}_{i_d,i_d}(\mathbf{k}, E)]. \tag{3-16b}$$

Lastly, an expression for the retarded defect self-energy follows directly from Eqs. (3-13a)-(3-13b) and (2-35)

$$\Sigma^R_d(E) = \frac{1}{2}[\Sigma^{>}_d(E) - \Sigma^{<}_d(E)] + \mathrm{i}\mathcal{P}\int_{-\infty}^{\infty} \frac{dE'}{2\pi} \frac{\Sigma^{>}_d(E) - \Sigma^{<}_d(E)}{E - E'}. \tag{3-17}$$

Notice that the self-energy expressions obtained in Eqs. (3-13a)-(3-13b) are identical to those shown in Eq. (13) of [65], where the author considers a single band model for both conduction and valence bands.

3.1.1 Detailed balance principle of the defect recombination rate

Current conservation requires the net recombination rate due to carrier capture and emission by the defect to vanish, as stated previously by Eq. (2-61). For a single defect at position $\mathbf{r}_d$, this condition can be restated in terms of the transition rate (2-64), becoming [65]

$$\bar{R} = \frac{1}{2\pi\hbar}\int dE\,[\Sigma^{<}_d(E)G^{>}_d(E) - \Sigma^{>}_d(E)G^{<}_d(E)] = 0, \tag{3-18}$$

which after expanding the defect self-energies leads to the condition

$$\bar{R}_{\mathrm{c}} = \bar{R}_{\mathrm{v}}, \tag{3-19}$$

with

$$\bar{R}_{\mathrm{c}} = \frac{+\mathrm{i}}{\hbar}\sum_{l\geq 0}\int dE\,\left[\mathcal{M}^{\mathrm{capt}}_d(l)\,n(z_d, E+\hbar\omega_l)G^{>}_d(E) + \mathcal{M}^{\mathrm{em}}_d(l)\,p(z_d, E+\hbar\omega_l)G^{<}_d(E)\right] \tag{3-20a}$$

$$\bar{R}_{\mathrm{v}} = \frac{-\mathrm{i}}{\hbar}\sum_{l\geq 0}\int dE\,\left[\mathcal{M}^{\mathrm{capt}}_d(l)\,p(z_d, E-\hbar\omega_l)G^{<}_d(E) + \mathcal{M}^{\mathrm{em}}_d(l)\,n(z_d, E-\hbar\omega_l)G^{>}_d(E)\right], \tag{3-20b}$$

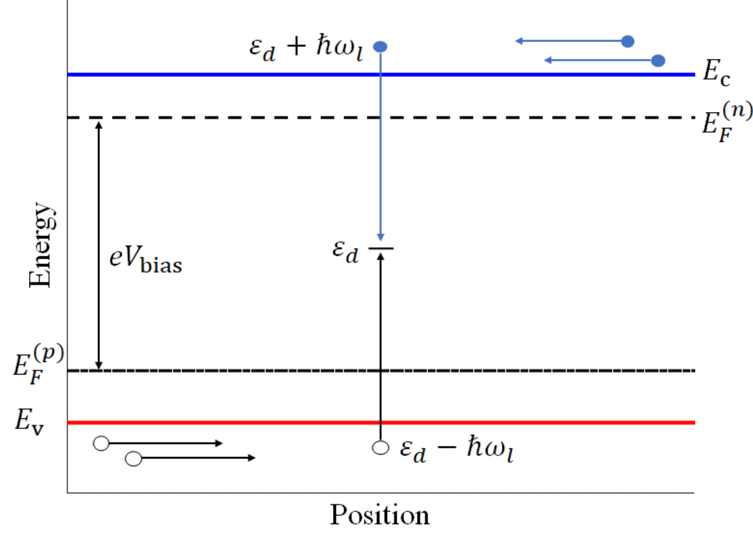


Figure 3-1: *Carrier recombination into a defect state due to multiphonon relaxation. At quasi-equilibrium conditions, electron and hole populations are characterized by Fermi statistics with quasi-Fermi levels $E_F^{(n)}$ and $E_F^{(p)}$, respectively. Electrons and holes coming from the right and left leads, respectively, recombine at the defect position through a multiphonon transition with energy $\hbar\omega_l$.*

the net electron and hole capture rates in the conduction and valence bands, respectively. The first term in Eq. (3-20a) can be interpreted as the rate at which electrons are captured from an energy $E + \hbar\omega_l$ in the conduction band into an energy E inside the gap, through multiphonon emission with energy $\hbar\omega_l$ (the matrix element $\mathcal{M}_d^{\text{capt}}$ contains the emission part of the self-energy). Conversely, the second term represents the rate at which electrons are emitted from a gap energy E into the conduction band energy $E + \hbar\omega_l$ via multiphonon absorption. A similar interpretation can be attributed to the net hole capture rate $\bar{R}_{\text{v}}$ in Eq. (3-20b). Fig. **3-1** shows an schematic representation of the electron and hole capture process into a defect state in a bulk semiconductor due to multiphonon relaxation.

The capture rates $\bar{R}_{\text{c}}$ and $\bar{R}_{\text{v}}$ have units s^{-1}, and therefore represent also the inverse lifetime of the defect capture process. If an ensemble of defects is considered instead, assuming that all are characterized by the same gap level ε_d, the volume capture rate can be recovered by multiplying the above rates by the defect density ρ_d

$$R_{\text{c(v)}} = \rho_d \, \bar{R}_{\text{c(v)}}. \tag{3-21}$$

However, the above expression is adding up the effect of an ensemble of isolated defects, but does not consider a potential carrier exchange between defects since our self-energy expression assumed only scattering between extended states and the confined state of a single defect. This imposes a limitation to the model for the case of a general defective material,

as the defects should be sufficiently far away from each other so they can be considered as independent (save for the indirect link via the electrostatics). In principle, inter-defect charge transfer is possible. For example, defect states may act as stepping stones in the *off* state of tunnel-field-effect transistors, while in the *on* state, they can affect the electron transport indirectly by modifying the device electrostatics [67, 68]. The possibility of multi-defect tunneling through a series of defects can be neglected if the distances are larger than the typical tunneling paths [69]. This condition is satisfied for defect densities lower than 10^{15}-10^{16} cm^{-3}, for which the average defect distance ranges between 50-100 nm. For such low defect densities, the trapped charge, which in principle may screen spontaneous and piezoelectric polarization fields, can be neglected in the self-consistent solution of Poisson's equation.

As was demonstrated in [65] using a single band model, Eq. (3-20a) leads to the semi-classical result for the electron capture rate obtained by Shockley and Read in [70], when quasi-equilibrium conditions are considered. Under such assumption, the carrier densities in the conduction band can be written as[3]

$$n(z_d, E) = f(E - E_F^{(n)}) D(z_d, E) \tag{3-22a}$$

$$p(z_d, E) = [1 - f(E - E_F^{(n)})] D(z_d, E), \tag{3-22b}$$

where $D(z_d, E)$ is the local density of states, f is the Fermi-Dirac distribution and $E_F^{(n)}$ is the quasi-Fermi level of the conduction band, and the equilibrium expressions (2-40) and (2-41) are assumed for the defect GFs

$$G_d^<(E) = \mathrm{i} f(E - E_F^{(d)}) A_d(E) \tag{3-23a}$$

$$G_d^>(E) = \mathrm{i}[f(E - E_F^{(d)}) - 1] A_d(E), \tag{3-23b}$$

where occupation is characterized by the Fermi-Dirac distribution with quasi-Fermi level $E_F^{(d)}$. Substituting these expressions in Eq. (3-21), the (semi-classical) volume capture rate is then written as

$$\begin{aligned} R_\mathrm{c}^\mathrm{eq} = \frac{\rho_d}{\hbar} \sum_{l \geq 0} \mathcal{M}_d^\mathrm{capt}(l) \int dE\, D(z_d, E + \hbar\omega_l) A_d(E) \\ \times \left\{ f(E + \hbar\omega_l - E_F^{(n)})[1 - f(E - E_F^{(d)})] - e^{-\beta\hbar\omega_l}[1 - f(E + \hbar\omega_l - E_F^{(n)})] f(E - E_F^{(d)}) \right\}, \end{aligned} \tag{3-24}$$

where we have used the relation (3-15) between the matrix elements for capture and emission.

[3]Notice that this two expressions follow from the definition of the LDOS in (2-84) and the quasi-equilibrium expressions of the GFs in (2-40) and (2-41).

The spectral function of the defect gives at equilibrium

$$\begin{aligned} A_{d,0}(E) &= \mathrm{i}[G^R_{d,0}(E) - G^A_{d,0}(E)] \\ &= \mathrm{i}\left[\frac{1}{E-\varepsilon_d - \mathrm{i}\eta} - \frac{1}{E-\varepsilon_d+\mathrm{i}\eta}\right] \\ &\overset{\eta\to 0}{=} 2\pi\,\delta(E-\varepsilon_d), \end{aligned} \tag{3-25}$$

which inserted in (3-24) cancels out the energy integration. Further using the relation

$$e^{-\beta\hbar\omega_l}[1 - f(E+\hbar\omega_l - E_F^{(n)})] = e^{\beta(E-E_F^{(n)})} f(E+\hbar\omega_l - E_F^{(n)}),$$

in (3-24), yields

$$\begin{aligned} R_{\mathrm{c}}^{\mathrm{eq}} &= \frac{2\pi\rho_d}{\hbar}\left\{[1-f(\varepsilon_d - E_F^{(d)})] - e^{\beta(\varepsilon_d - E_F^{(n)})} f(\varepsilon_d - E_F^{(d)})\right\} \\ &\times \sum_{l\geq 0} \mathcal{M}_d^{\mathrm{capt}}(l)\, D(z_d, \varepsilon_d + \hbar\omega_l) f(\varepsilon_d + \hbar\omega_l - E_F^{(n)}) \\ &= \frac{2\pi\rho_d}{\hbar}\big[1 - f(\varepsilon_d - E_F^{(d)})\big]\big[1 - e^{\beta(E_F^{(d)} - E_F^{(n)})}\big] \\ &\times \sum_{l\geq 0} \mathcal{M}_d^{\mathrm{capt}}(l)\, D(z_d, \varepsilon_d + \hbar\omega_l) f(\varepsilon_d + \hbar\omega_l - E_F^{(n)}) \\ &= \rho_d\big[1 - f(\varepsilon_d - E_F^{(d)})\big]\big[1 - e^{\beta(E_F^{(d)} - E_F^{(n)})}\big] \int dE\, c_n(E) D(z_d, E) f(E - E_F^{(n)}), \end{aligned} \tag{3-26}$$

where we have introduced the quantity

$$c_n(E) = \frac{2\pi}{\hbar}\sum_{l\geq 0} \mathcal{M}_d^{\mathrm{capt}}(l)\delta(E - \varepsilon_d - \hbar\omega_l). \tag{3-27}$$

Eq. (3-26) shows the same result as the one obtained by Shockley and Read in Eq. (2.10) of [70], and the quantity c_n is called the *electron capture coefficient*, in units $\mathrm{m}^3\mathrm{s}^{-1}$. The previous result is interesting as it establishes a connection between the NEGF capture matrix element $\mathcal{M}_d^{\mathrm{capt}}$ and the semi-classical capture coefficient c_n. Notice also that at equilibrium, *i.e.*, when $E_F^{(n)} = E_F^{(d)} = E_F$, the net electron capture rate (3-26) becomes zero (because of the exponential term with the difference of the quasi-Fermi levels), showing that both capture and emission processes of electrons from and to the conduction band become perfectly balanced. A similar result can be obtained for the net hole capture rate R_{v}.

3.1.2 Derivation of the multiphonon matrix elements

The capture coupling element, related with the square of the absolute value of the electron-phonon matrix element, is the only ingredient missing for the computation of the defect

self-energies, as the emission term follows directly from (3-15). A first approximation of this term can be obtained by substituting the electron-phonon potential in (3-14a), using an optical deformation potential model (see [71] and [59], Sec. 3.4)

$$|M(l,\mathbf{q})|^2 = \frac{\hbar D_{tK}^2}{2V\rho\,\omega_l}, \tag{3-28}$$

where D_{tK} is the optical deformation potential, ρ is the (mass) density of the material and $\hbar\omega_l$ is the multiphonon energy. The *deformation potential capture matrix element* is then given by

$$\mathcal{M}_{d,\mathrm{def}}^{\mathrm{capt}}(l) = \frac{\Omega_d}{\Omega}(n_l+1)\frac{\hbar D_{tK}^2}{2\rho\omega_l}, \tag{3-29}$$

with Ω the unit cell volume. Likewise, the polar optical potential (2-157) can be used to determine the *polar optical capture matrix element*, yielding

$$\begin{aligned}\mathcal{M}_{d,\mathrm{pop}}^{\mathrm{capt}}(l) &= \Omega_d\,(n_l+1)\sum_{\mathbf{q}}\frac{e^2\hbar\omega_{\mathrm{LO}}}{2V}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)\left(\frac{q}{q^2+q_0^2}\right)^2 \\ &\approx \Omega_d\,(n_l+1)\frac{e^2\hbar\omega_l}{2}\left(\frac{1}{\epsilon_\infty}-\frac{1}{\epsilon_s}\right)\frac{1}{2\pi^2}\int_0^{\pi/a_L} dq\,\left(\frac{q^2}{q^2+q_0^2}\right)^2.\end{aligned} \tag{3-30}$$

On the other hand, in Ref. [65] it is presented another method to find the capture matrix element that directly links this value with the electron capture rate from multiphonon relaxation theory in [66]. The starting point is the relation (3-27) between the capture matrix element and the electron capture coefficient, which after a direct comparison with the multiphonon electron and hole capture coefficients from Eq. (12) in [66], lead to the band dependent capture matrix elements[4]

$$\mathcal{M}_{\mathrm{dc}}^{\mathrm{capt}}(l) = \mathcal{M}_{\mathrm{dc}}^{0}\,\frac{(S-l)^2}{S}\,L(l) \tag{3-31a}$$

$$\mathcal{M}_{\mathrm{dv}}^{\mathrm{capt}}(l) = \mathcal{M}_{\mathrm{dv}}^{0}\,\frac{(S+l)^2}{S}\,L(l), \tag{3-31b}$$

where S is the Huang-Rhys factor, and $\mathcal{M}_{\mathrm{dc(v)}}^{0} = \hbar^2 c_{n(p)}^0 r_{\mathrm{ph},n(p)}^2$ are unknown coefficients containing the electron-phonon (hole-phonon) squared matrix element $r_{\mathrm{ph},n(p)}^2$ and the prefactors $c_{n(p)}^0$ which include various unknown quantities like the impurity potential strength, the symmetry of the wave functions and the localization radius, etc. The function $L(l)$ is the lineshape function of multiphonon theory

$$L(l) = e^{-S(2n_0+1)}\left(\frac{n_0+1}{n_0}\right)^{l/2} I_l(2S\sqrt{n_0(n_0+1)}), \tag{3-32}$$

[4] In writing this result we neglect the electron-filed matrix element appearing in the multi-phonon electron capture coefficient in [66], as it is suggested by the author.

with $n_0 = \left(e^{-\beta\hbar\omega_0} - 1\right)$ the (single) phonon occupation number and I_l the modified Bessel function. The coefficients $\mathcal{M}^0_{\mathrm{dc(v)}}$ are determined through the definition of the (bulk) average capture coefficients

$$\langle c_n \rangle = \frac{1}{N_\mathrm{C}} \int_{E_\mathrm{c}}^{\infty} dE\, e^{\beta(E_\mathrm{c}-E)} c_n(E) D(z_d, E) \tag{3-33a}$$

$$\langle c_p \rangle = \frac{1}{N_\mathrm{V}} \int_{-\infty}^{E_\mathrm{v}} dE\, e^{\beta(E-E_\mathrm{v})} c_p(E) D(z_d, E), \tag{3-33b}$$

and their relation with the (zero-field) SRH carrier lifetimes $\tau^{-1}_{n(p),0} = \rho_d \langle c_{n(p)} \rangle$, leading to the expressions

$$\mathcal{M}^0_{\mathrm{dc}} = \frac{\hbar\, N_\mathrm{C}}{2\pi\, \tau_{n,0}\rho_d} \left[\sum_{l\geq 0} \frac{(S-l)^2}{S} L(l)\, e^{\beta(E_\mathrm{c}-\varepsilon_d-\hbar\omega_l)} D(z_d, \varepsilon_d + \hbar\omega_l) \right]^{-1} \tag{3-34a}$$

$$\mathcal{M}^0_{\mathrm{dv}} = \frac{\hbar\, N_\mathrm{V}}{2\pi\, \tau_{p,0}\rho_d} \left[\sum_{l\geq 0} \frac{(S+l)^2}{S} L(l)\, e^{\beta(\varepsilon_d-\hbar\omega_l-E_\mathrm{v})} D(z_d, \varepsilon_d - \hbar\omega_l) \right]^{-1} . \tag{3-34b}$$

In terms of the band dependent multiphonon capture and emission matrix elements, Eqs. (3-31a) and (3-31b), respectively, the defect self-energies are now given by

$$\Sigma^<_d(E) = +2\pi\mathrm{i} \sum_{l\geq 0} \left[\mathcal{M}^{\mathrm{capt}}_{\mathrm{dc}}(l)\, n(z_d, E+\hbar\omega_l) + \mathcal{M}^{\mathrm{em}}_{\mathrm{dv}}(l)\, n(z_d, E-\hbar\omega_l) \right] \tag{3-35a}$$

$$\Sigma^>_d(E) = -2\pi\mathrm{i} \sum_{l\geq 0} \left[\mathcal{M}^{\mathrm{capt}}_{\mathrm{dv}}(l)\, p(z_d, E-\hbar\omega_l) + \mathcal{M}^{\mathrm{em}}_{\mathrm{dc}}(l)\, p(z_d, E+\hbar\omega_l) \right] . \tag{3-35b}$$

Fig. **3-2** shows a comparison of the NEGF volume recombination rate from Eq. (3-21) for bulk GaN, with the capture matrix element computed using an optical deformation potential model (3-29), a polar optical potential model (3-30) and multiphonon theory (3-31a)-(3-31b). For the optical deformation potential model, $D_{tK} = 3 \times 10^9$ eVcm^{-1} was used and $\Omega_d/\Omega \approx 1$ was assumed. On the other hand, the multi-phonon theory approach used as carrier lifetimes $\tau_{n,0} = 9$ ns and $\tau_{p,0} = 46$ ns. The semiclassical result obtained from the SRH formula[5]

$$R_{SRH} = \frac{np - n_i^2}{(n+n_1)\tau_{p,0} + (p+p_1)\tau_{n,0}}, \tag{3-36}$$

is also shown for comparison, where n_i is the intrinsic carrier density of the material, and n_1 and p_1 are the equivalent electron and hole densities when the Fermi level coincides with the trap level, respectively[6]. The recombination rate obtained from multiphonon theory is able

[5] Notice that Eqs. (3-26) and (3-36) are equivalent formulations of the defect recombination rate in Bulk semiconductors. See Ref. [70], Eq. (4.4).

[6] The expressions for these quantities can be found in [70], Eq. (3.7).

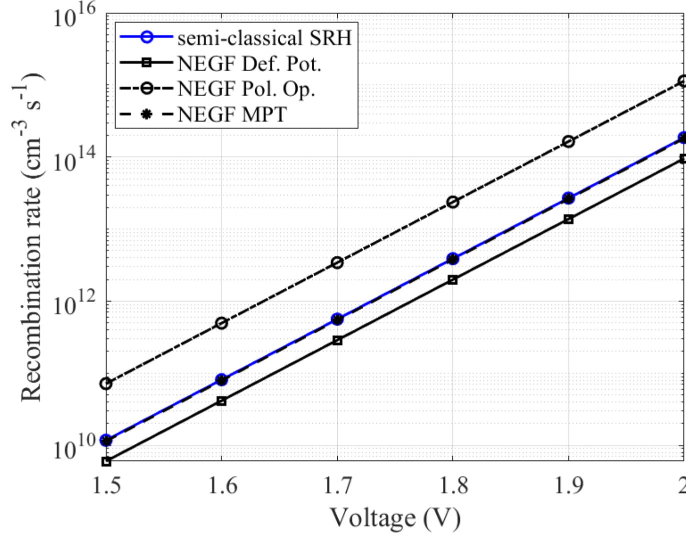


Figure 3-2: *Comparison of the defect recombination rate for bulk GaN obtained from the semi-classical formula (3-36) and the NEGF rate (3-21) by using different models for the capture matrix element: optical deformation potential* (3-29)*, polar optical potential* (3-30) *and multiphonon theory (MPT)* (3-31a)-(3-31b)*. A defect density $\rho_d = 10^{15}$ cm^{-3} was used in all the computations.*

to reproduce very closely the semi-classical results as it depends on the same parameters, *i.e.*, the carrier lifetimes, and the coefficients (3-34a)-(3-34b) ensure that the average semiclassical capture coefficient is maintained. It is also interesting to observe that the three models (*i.e.*, the net capture rate coming from the deformation potential, polar optical, and multiphonon theory models) are able to reproduce the ideality factor $\eta = 2$, typical of the zero-field Shockley-Read-Hall recombination process.

3.2 Shockley-Read-Hall self-energy

In the previous section, phonon interaction with deep single-level defects was investigated. Such interaction introduces a renormalization in the defect GF, that reflects a change in the defect occupation due to scattering of carriers from and to extended states in the device. The task is now to determine the effect of defect scattering in the device GF, for which the corresponding self-energy should be determined. Since our NEGF framework for layered semiconductors assumes the material structure to be homogeneous in the transversal direction, point defects (which would introduce inhomogeneities in the material) have to be taken into account indirectly by means of a scattering self-energy. Let us assume an ensemble of uncoupled, strongly localized point defects participating in nonradiative recombination

events through phonon interactions. From the second-order contribution of the perturbation expansion of the GFs in Eq. (3-6), the self-energy can be identified as

$$\Sigma(12) = \langle \hat{T}_C \{V_{ep}(1)V_{ep}(2)\}\rangle\, G(12). \tag{3-37}$$

Assuming that the self-energy acts locally at each defect position, we approximate the full propagator $G(12)$ with that of the defect ensemble

$$G(\mathbf{r},t;\mathbf{r}',t') = \sum_{i_d} \Omega_d\, \delta(\mathbf{r}-\mathbf{r}_{i_d})\, \delta(\mathbf{r}'-\mathbf{r}_{i_d})\, G_d(t,t'), \tag{3-38}$$

where i_d labels each isolated defect position $\mathbf{r}_{i_d}$, characterized the single level energy ε_d, with Ω_d the localization volume and G_d the defect GF. The covariant representation of the self-energy (2-68) in the discrete basis (2-77) is given by

$$\begin{aligned}
\Sigma_{im,jn}(\mathbf{k};t,t') &= \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{im,\mathbf{k}}(\mathbf{r})\, \Sigma(\mathbf{r},t;\mathbf{r}',t')\, \phi_{jn,\mathbf{k}}(\mathbf{r}') \\
&= \int d\mathbf{r} \int d\mathbf{r}'\, \phi^*_{im,\mathbf{k}}(\mathbf{r})\, \langle \hat{T}_C \{V_{ep}(\mathbf{r},t)V_{ep}(\mathbf{r}',t')\}\rangle \\
&\times \left[\sum_{i_d} \Omega_d\, \delta(\mathbf{r}-\mathbf{r}_{i_d})\, \delta(\mathbf{r}'-\mathbf{r}_{i_d})\, G_d(t,t')\right] \phi_{jn,\mathbf{k}}(\mathbf{r}') \\
&= \frac{\mathrm{i}\hbar}{A} \sum_{l\geq 0} \sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 D_0(l,\mathbf{q};t,t') \\
&\times \sum_{i_d} u^*_{m,0}(\mathbf{r}_{i_d}) t_i(z_{i_d})\, \Omega_d\, G_d(t,t')\, u_{n,0}(\mathbf{r}_{i_d}) t_j(z_{i_d}),
\end{aligned} \tag{3-39}$$

where Eq. (3-10) has been used. At this point, the Bloch function products are approximated by lattice cell averages [30], and the sum over defect positions is partitioned over each material slab $l_e = [z_e, z_{e+1}]$

$$\sum_{i_d} \equiv \sum_e \sum_{\mathbf{r}_{i_d}\in V_e} .$$

where V_e is the slab volume. Moreover, we assume the defect density to be constant over each slab, and that defects belonging to the same slab are characterized by the same confined level energy. Similar to the approximation used for reciprocal space sums, the sum over defect positions is approximated as

$$\sum_{\mathbf{r}_{i_d}\in V_e} \approx \left(\frac{V_e}{N_d}\right)^{-1} \int_{V_e} d\mathbf{r} = \rho_d \int_{V_e} d\mathbf{r},$$

where N_d is the total number of defects inside the slab l_e, and $\rho_d = N_d/V_e$ is the average defect density on the slab l_e. With the previous considerations, the SRH self-energy reads

$$\Sigma_{im,jn}(\mathbf{k};t,t') = \mathrm{i}\hbar \sum_{l\geq 0}\sum_{\mathbf{q}} |M(l,\mathbf{q})|^2 D_0(l,\mathbf{q};t,t')\,\delta_{m,n} \times \sum_e \Omega_d\, G_d(t,t')\,\rho_d(z_e)\int_{l_e} dz\, t_i(z)t_j(z). \tag{3-40}$$

As usual, we proceed to apply the Langreth rules in order to obtain the real-time (lesser and greater) components of the self-energy, and Fourier transform with respect to the time difference $\tau = t-t'$, which results in the steady-state expressions of the self-energy. Assuming an equilibrium phonon population with constant phonon energy $\hbar\omega_0$, and with $\hbar\omega_l$ the energy of a multiphonon transition involving l phonons, the final expression of the SRH self-energy is given by

$$\Sigma^{\lessgtr}_{im,jn}(\mathbf{k},E) = \delta_{m,n}\sum_{l\geq 0}\left[\mathcal{M}_d^{\mathrm{capt}}(l)G^{\lessgtr}_{d,ij}(E\pm\hbar\omega_l) + \mathcal{M}_d^{\mathrm{em}}(l)G^{\lessgtr}_{d,ij}(E\mp\hbar\omega_l)\right], \tag{3-41}$$

with the multiphonon matrix elements $\mathcal{M}_d^{\mathrm{em/capt}}$ given by Eqs. (3-14a)-(3-14b), and

$$G^{\lessgtr}_{d,ij}(E) = \sum_e G^{\lessgtr}_d(E)\rho_d(z_e)\int_{l_e} dz\, t_i(z)t_j(z) \approx G^{\lessgtr}_d(E)\rho_d(z_i)M_{i,j}, \tag{3-42}$$

where the defect GFs $G^{\lessgtr}_d$, given by Eqs. (3-4a)-(3-4b), depend on the longitudinal position (and therefore on the element l_e) through the defect level $\varepsilon_d \equiv \varepsilon_d(z)$[7] and the defect self-energies, which depend on the local carrier densities as shown in Eqs. (3-13a)-(3-13b). Each time the SRH self-energies in (3-41) are computed, the defect GFs $G^{\lessgtr}_d$ have to be computed self-consistently by solving the Dyson's and Keldysh equations (3-4a)-(3-4b) and including the defect self-energies in Eqs. (3-13a)-(3-13b). For the case of deep defects, since the defect GF is assumed to be negligible far away from the defect energy ε_d, one of the two terms in (3-41) can be neglected depending on the band to which the energy E belongs, yielding the approximation

$$\Sigma^{\lessgtr}_{im,jn}(\mathbf{k},E) = \delta_{m,n}\sum_{l\geq 0}\mathcal{M}_d^{\mathrm{em/capt}}(l)G^{\lessgtr}_{d,ij}(E-\hbar\omega_l) \quad \text{if} \quad E\gg\varepsilon_d \tag{3-43a}$$

$$\Sigma^{\lessgtr}_{im,jn}(\mathbf{k},E) = \delta_{m,n}\sum_{l\geq 0}\mathcal{M}_d^{\mathrm{capt/em}}(l)G^{\lessgtr}_{d,ij}(E+\hbar\omega_l) \quad \text{if} \quad E\ll\varepsilon_d. \tag{3-43b}$$

Analogous to the case of carrier-phonon self-energy, the expression of the SRH self-energies can be further simplified by assuming that the defect GF remains unperturbed by the interaction with carriers in the extended states, so the quasi-equilibrium expressions (3-23)

[7] Notice that the defect energy level should be adjusted with the Hartree potential producing the band bending in the device, *i.e.*, $\varepsilon_d(z) \equiv \varepsilon_{d,0}(z) + U_{\mathrm{Hartree}}(z)$ [65], where $\varepsilon_{d,0}(z)$ is the flat-band energetic position of the defect.

can be used. This reduces the computation of the defect GFs to finding the corresponding quasi-Fermi level $E_F^{(d)}$ such that the condition (3-18) for zero net recombination current is satisfied, yielding

$$-\frac{1}{2\pi\hbar}\int dE\left\{\mathrm{i}\Sigma_d^<(E)\left[1-f(E-E_F^{(d)})\right]A_d(E)+\mathrm{i}\Sigma_d^>(E)f(E-E_F^{(d)})A_d(E)\right\}=0$$

$$\int dE\left\{\mathrm{i}\Sigma_d^<(E)A_d(E)+\Gamma_d(E)f(E-E_F^{(d)})A_d(E)\right\}=0, \quad (3\text{-}44)$$

with $\Gamma_d(E) = i\,[\Sigma_d^>(E) - \Sigma_d^<(E)]$ the defect broadening function. Inserting the equilibrium expression of the defect spectral function (3-25) in (3-44), leads to the result

$$\Sigma_d^<(\varepsilon_d) = \mathrm{i}f(\varepsilon_d - E_F^{(d)})\Gamma_d(\varepsilon_d), \quad (3\text{-}45)$$

from which the occupation function follows directly. The result in Eq. (3-45) is not surprising as it is the quasi-equilibrium relation between the defect self-energy and broadening function (see Eq. (2-42)). Although with this approximation the self-consistent solution of Dyson's and Keldysh equations for the defect GF is avoided, it is still necessary to compute the defect self-energies as the occupation function depends on them through Eq. (3-45). Since the spectral function is given by a delta in energy (although some artificial broadening can be included by using a small η), the sum over phonon modes is canceled out leaving the single mode l_0 such that $\varepsilon_d = E \mp l_0\hbar\omega_0$. Numerically, this delta in energy is approximated by the square function $\delta(E-\varepsilon_d) \approx 1/\Delta E$, for $E \in [\varepsilon_d - \Delta E/2, \varepsilon_d + \Delta E/2]$, with ΔE the mesh spacing in the energy grid. In our subsequent simulations, we will use this last approximation for the SRH self-energy, with the capture matrix elements based on multiphonon theory from Eqs. (3-31a)-(3-31b). Since this terms are band dependent, $\mathcal{M}_d^{\mathrm{em/capt}}$ in Eqs. (3-43a)-(3-43b) should be replaced by $\mathcal{M}_{\mathrm{dc}}^{\mathrm{em/capt}}$ for $E \gg \varepsilon_d$ and by $\mathcal{M}_{\mathrm{dv}}^{\mathrm{em/capt}}$ for $E \ll \varepsilon_d$.

3.3 Enhancement of the defect recombination by tunneling

Defect recombination in layered semiconductors may deviate drastically from the bulk quasi-equilibrium situation described by Eq. (3-26). By introducing the defect thermal binding energy $E_t = E_\mathrm{c} - \varepsilon_d$, measured with respect to the conduction band edge, the defect lesser self-energy (3-35a), which may be interpreted as an in-scattering rate of electrons coming from the extended states, can be split into the contribution of two terms[8]

$$\Sigma_d^<(E) = \Sigma_{d,\mathrm{Direct}}^<(E) + \Sigma_{d,\mathrm{TAT}}^<(E) \quad (3\text{-}46)$$

[8]The greater self-energy, which can be interpreted as an in-scattering of holes coming from the extended states, can be split into direct and TAT contributions in an analogous way.

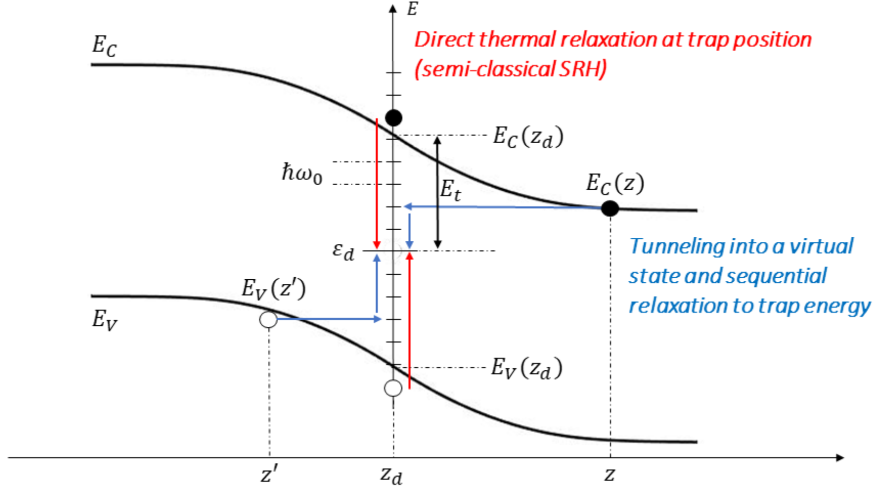


Figure 3-3: *Schematic representation of the recombination of carriers through a defect state at z_d. Electron and hole capture into the defect level may proceed in two different ways: by direct thermal relaxation at z_d of electrons above $E_c(z_d)$ (holes below $E_v(z_d)$), or through tunneling of an electron from a position z (hole from a position z') into a sub-gap state at z_d and sequential relaxation to the defect energy.*

with

$$\Sigma^<_{d,\text{Direct}}(E) = 2\pi\mathrm{i} \sum_{l\geq\frac{E_t}{\hbar\omega_0}} \mathcal{M}^{\text{capt}}_{\text{dc}}(l)\, n(z_d, E+\hbar\omega_l) + 2\pi\mathrm{i} \sum_{l\geq\frac{E_g-E_t}{\hbar\omega_0}} \mathcal{M}^{\text{em}}_{\text{dv}}(l)\, n(z_d, E-\hbar\omega_l) \tag{3-47a}$$

$$\Sigma^<_{d,\text{TAT}}(E) = 2\pi\mathrm{i} \sum_{\frac{E_t}{\hbar\omega_0}>l\geq 0} \mathcal{M}^{\text{capt}}_{\text{dc}}(l)\, n(z_d, E+\hbar\omega_l) + 2\pi\mathrm{i} \sum_{\frac{E_g-E_t}{\hbar\omega_0}>l\geq 0} \mathcal{M}^{\text{em}}_{\text{dv}}(l)\, n(z_d, E-\hbar\omega_l), \tag{3-47b}$$

representing electron capture from energies beyond the band edges ($\Sigma^<_{d,\text{Direct}}$) and inside the gap ($\Sigma^<_{d,\text{TAT}}$). Within the semiclassical treatment of SRH, the spectral carrier densities are assumed to be zero for energies in the gap so $\Sigma^<_{d,\text{TAT}}(E) = 0$, and carrier capture is assumed to be through direct thermal relaxation at the trap position, *i.e.*, all the contribution comes from $\Sigma^<_{d,\text{Direct}}(E)$. This assumption is true in bulk semiconductors at flat-band conditions (zero electric field), where SRH recombination is characterized by an ideality factor of two at low voltages in the I-V curve. However, $\Sigma^<_{d,\text{TAT}}(E)$ does not vanish in general, because of the presence of *tail states* that split from the bands into the gap, and which enable *trap-assisted*

tunneling (TAT), *i.e.*, a two-step capture process in which electrons tunnel into a sub-gap state and are subsequently captured by multiphonon emission/absorption. Fig. **3-3** shows an schematic representation of these two situations, where the capture of carriers via a deep defect level may proceed by direct thermal relaxation or through TAT.

In addition, the SRH self-energies from Eqs. (3-43a)-(3-43b) can also be divided into direct and TAT contributions depending if the energy is above or below the band edges, which can be expressed equivalently by splitting the sum over phonon modes, obtaining

$$\Sigma^{\lessgtr}_{im,jn}(\mathbf{k},E) = \underbrace{\delta_{m,n} \sum_{l \geq \frac{E_t}{\hbar\omega_0}} \mathcal{M}^{\text{em/capt}}_{\text{dc}}(l) G^{\lessgtr}_{d,ij}(E - \hbar\omega_l)}_{\text{Direct}} + \underbrace{\delta_{m,n} \sum_{\frac{E_t}{\hbar\omega_0} > l \geq 0} \mathcal{M}^{\text{em/capt}}_{\text{dc}}(l) G^{\lessgtr}_{d,ij}(E - \hbar\omega_l)}_{\text{TAT}} \quad \text{if} \quad E \gg \varepsilon_d \tag{3-48a}$$

$$\Sigma^{\lessgtr}_{im,jn}(\mathbf{k},E) = \underbrace{\delta_{m,n} \sum_{l \geq \frac{E_g - E_t}{\hbar\omega_0}} \mathcal{M}^{\text{capt/em}}_{\text{dv}}(l) G^{\lessgtr}_{d,ij}(E + \hbar\omega_l)}_{\text{Direct}} + \underbrace{\delta_{m,n} \sum_{\frac{E_g - E_t}{\hbar\omega_0} > l \geq 0} \mathcal{M}^{\text{capt/em}}_{\text{dv}}(l) G^{\lessgtr}_{d,ij}(E + \hbar\omega_l)}_{\text{TAT}} \quad \text{if} \quad E \ll \varepsilon_d. \tag{3-48b}$$

Let us consider a GaN *p*-*n* diode, composed of two highly doped layers with dopant concentrations $N_A = N_D = 10^{19}$ cm^{-3}. The device structure with its respective measures is shown in Fig. **3-4**, and the material parameters can be found in Tab. **3-1**. The enhancement of the defect recombination due to tail state formation can be studied using our NEGF implementation on this simple test structure, where the high doping densities create a strong electric field close to the junction, which favors the tunneling of carriers into subgap states.

Table 3-1: Parameters used in the simulation of the GaN diode.

	Parameter	GaN	Unit
Bandgap energy	E_g	3.44	eV
Electron effective mass	m_e	0.2	m_0
Hole effective mass	m_h	1.89	m_0
static dielectric constant	ϵ	9.5	ϵ_0
High frequency dielectric constant	ϵ_∞	5.35	ϵ_0
Polar optical phonon energy	$\hbar\omega_{\text{LO}}$	90	meV
Acoustic deformation potential	D_a	8	eV

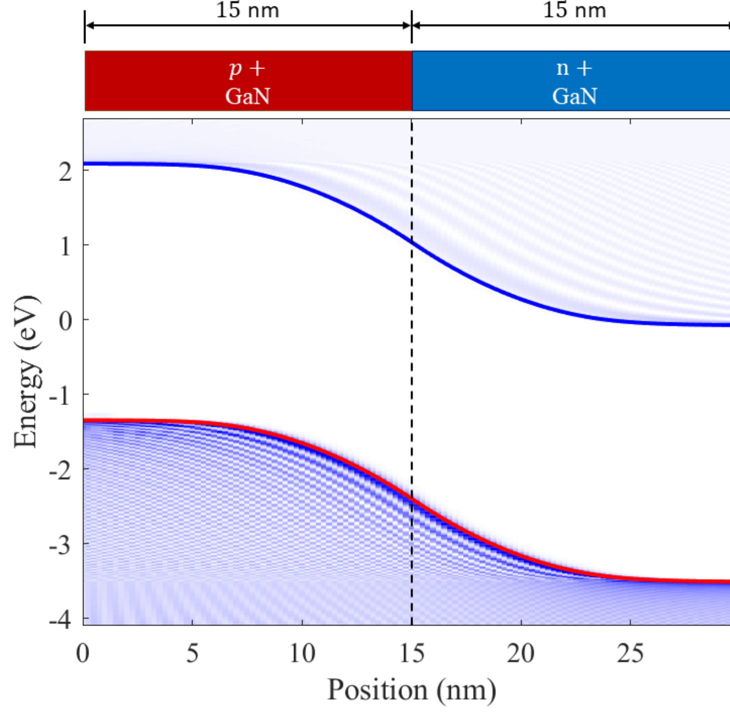


Figure 3-4: *Schematic of the p-n diode studied in this section, with the corresponding physical dimensions of each layer. The local density of states at* $\boldsymbol{k} = 0$ *can be observed, showing the typical interference pattern due to the coherent superposition of the wavefunctions for carriers injected from the contacts and reflected at the junction.*

Band tail formation can be observed in Fig. 3.5(a), where diode was simulated using an 8-band $\mathbf{k} \cdot \mathbf{p}$ model (the band parameters can be found in Tab. **A-2**), including acoustic and polar optical phonon scattering and with an applied bias voltage $V_{\text{bias}} = 1.3$ V. The right-hand side of the figure shows a cut of the local density of states at the junction position $z_0 = 15$ nm where the electric field is maximum, with the horizontal dashed lines marking the band edges. The density of states shows exponentially decaying tails that extend deep into the gap, which are the result of the band bending induced by the electric field. Further increasing the bias voltage to $V_{\text{bias}} = 3.4$ V, when the band diagram is close to flat-band condition and the field-induced tails become negligible, Fig. 3.5(b) shows the formation of band tails induced by phonon interactions, with exponentially decreasing LO-phonon resonances extending into the gap.

Since most of the recombination occurs in high-electric-field regions where the contribution of carrier-phonon interaction to the formation of tail states is negligible, most semiclassical models attempting to represent trap-assisted tunneling take into account only the effect of the electric field in the local density of states. Among such models can be found the model from *Hurkx et al.* [72] and the *Schenk* model [73], in which the contribution of the field-enhanced

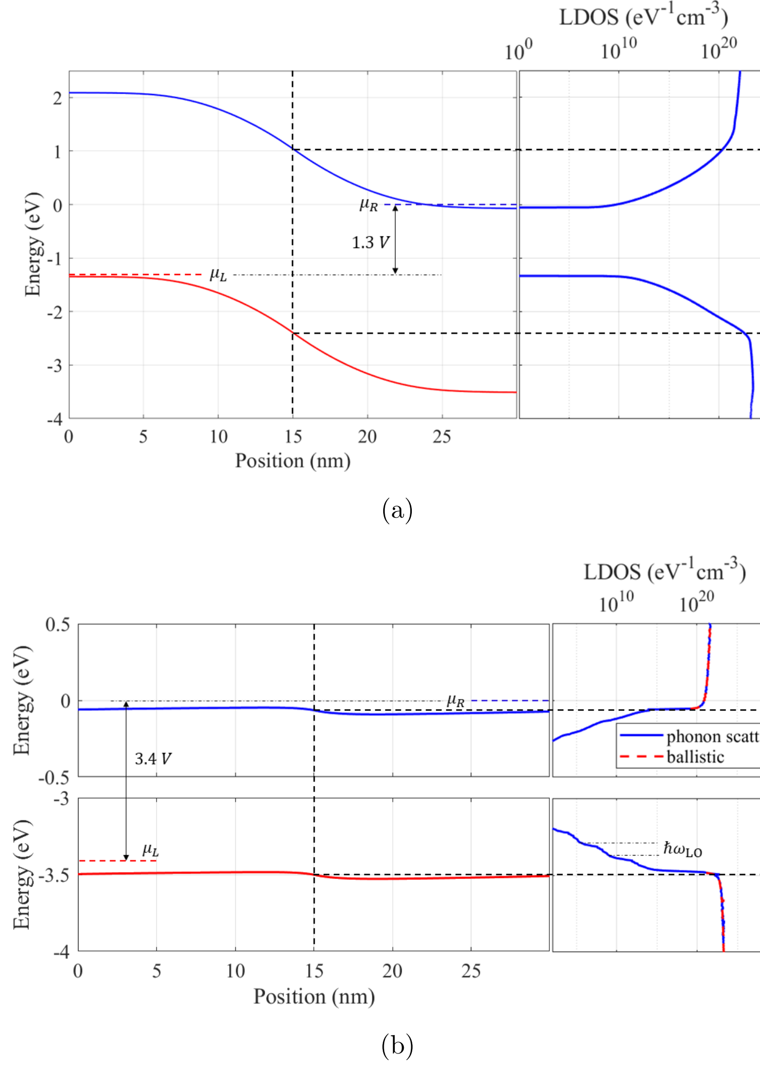


Figure 3-5: *Band diagram of the p-n junction at (a) $V_{bias} = 1.3$ V and (b) $V_{bias} = 3.4$ V. A cut of the LDOS at $z_0 = 15$ nm is shown on the right-hand side of each plot. At low forward bias voltage, the high built-in electric field generate tail states that extend deep inside the gap. At high voltages, when the material is close to flat-band condition and the internal fields are negligible, the tails show LO-phonon resonances.*

tail states is found analytically and the usual SRH (zero-field) lifetimes $\tau_{n(p),0}$ are replaced by the field-dependent lifetimes $\tau_{n(p)}(F) = \tau_{n(p),0}\, g_{n(p)}[F(z)]$, where $g_{n(p)}[F(z)]$ are the field-enhancement factors which depend on the local values of the electric field $F(z)$. Although relatively simple in its application, these models have the disadvantage of increasing the number of unknown parameters by requiring additional quantities like the tunneling masses, which are often used as fitting parameters in device simulations. Along with the field and phonon induced tails, subgap states may also result from other scattering processes (carrier-

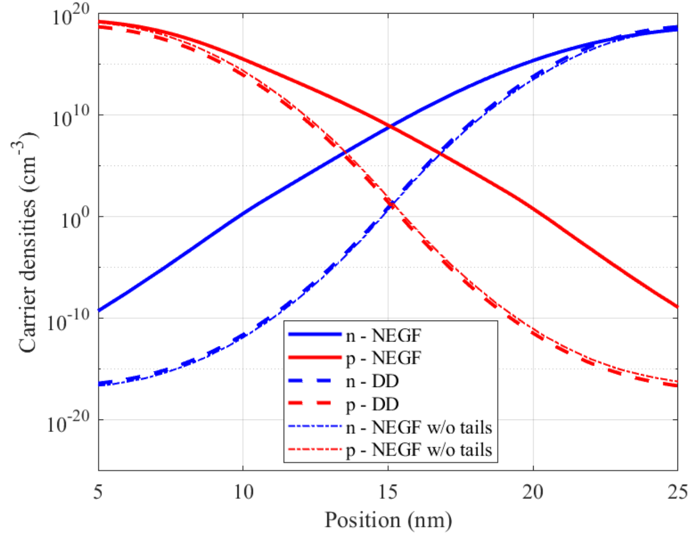


Figure 3-6: *Comparison of the NEGF and drift-diffusion (DD) carrier densities in the p-n device at* $V_{bias} = 1.3$ *V. When the field-induced tail contribution is excluded from the carrier densities computation, both NEGF and DD densities coincide. Minority carrier population is greatly enhanced on each side of the junction due to the high electric field. Far from the junction the electric field becomes negligible and both NEGF and DD densities converge to similar values.*

carrier, impurity, etc) or due to confinement effects where bound states leak into the barriers. The interested reader can find in [74] a detailed discussion on band-tail formation and band-gap narrowing due to polar optical and impurity scattering using a NEGF framework.

The increment in the electron and hole populations due to the field-induced tails can be observed by comparing the quantum charge densities obtained from NEGF, with the semi-classical charge obtained, *e.g.*, from a drift-diffusion simulation. In this opportunity, the NEGF simulation was performed using a two-band effective mass model (see the Hamiltonian in Eq. (2-102)) including the first conduction and heavy-hole bands, with the effective masses shown in Tab. **3-1**, and the same masses were used in drift-diffusion. Fig. **3-6** shows the carrier densities at $V_{\text{bias}} = 1.3$ V, computed with NEGF and drift-diffusion, where it can be seen that the carrier densities are greatly increased due to the additional states available by virtue of the field-induced tails. This effect is particularly evident in the increase of the minority carrier population on each side of the junction, while far from the junction both NEGF and drift-diffusion carrier densities converge to similar values as it is expected in neutral regions (*i.e.*, where the electric field is zero). This figure displays also the NEGF carrier densities without the contribution of the tail states, *i.e.*, they are computed through

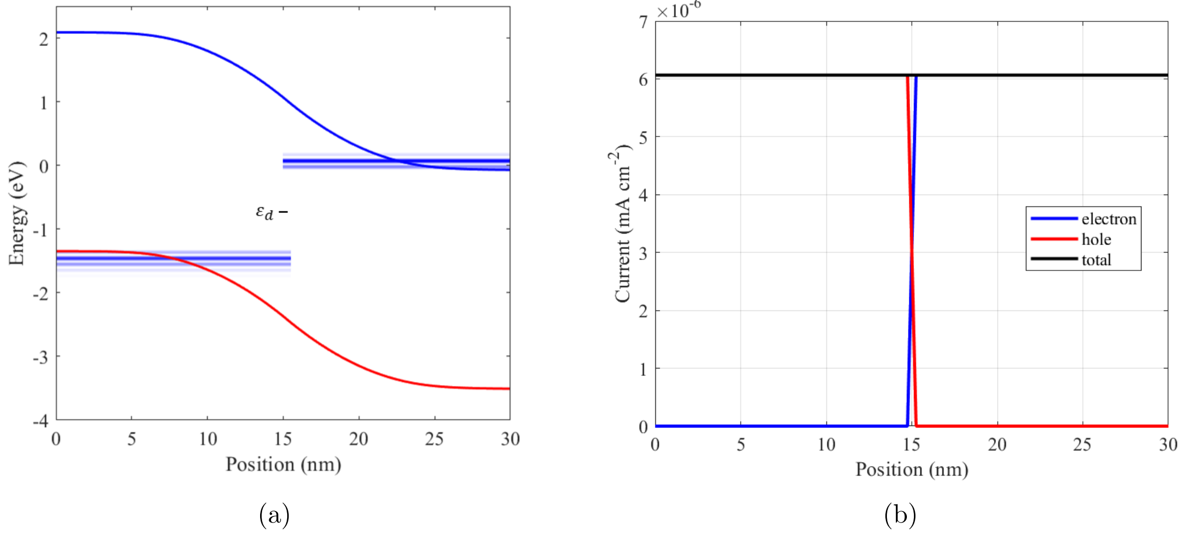


Figure 3-7: *(a) Spectrally resolved current density at $V_{bias} = 1.3$ V. Carriers coming from the contacts tunnel to the defect at $z_d = 15$ nm, and recombine through multi-phonon relaxation. The continuous stripes, which represent current carrying channels, are a signature of the multi-phonon recombination process as it can only happen at integer multiples of the effective phonon energy $\hbar\omega_0$ away from ε_d. (b) The current density, obtained from the integration of the spectral current density over energy, shows that the SRH self-energy satisfy the conservation of the total current.*

the following expressions

$$n(z_i) = -\mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int_{E_\mathrm{c}(z_i)}^{\infty}\frac{dE}{2\pi}\mathrm{Tr}[\mathbf{G}^{<}_{i,i}(\mathbf{k},E)] \tag{3-49a}$$

$$p(z_i) = \mathrm{i}\frac{1}{A}\sum_{\mathbf{k}}\int_{-\infty}^{E_\mathrm{v}(z_i)}\frac{dE}{2\pi}\mathrm{Tr}[\mathbf{G}^{>}_{i,i}(\mathbf{k},E)], \tag{3-49b}$$

where the energy integration is restricted to the sharp band edges. It is interesting to observe that the resulting carrier densities computed in this way using the NEGF results coincide with the semiclassical densities from the drift-diffusion simulation. This is true because the NEGF LDOS in this example resembles the square root behavior of bulk materials beyond the band edges, although some differences may be expected in different devices featuring heterojunctions and quantum wells.

We proceed further in our analysis by introducing a single defect at position $z_d = 15$ nm, where the electric field is maximum, with the defect level located at mid-gap $\varepsilon_d = E_\mathrm{c}(z_d) - E_g/2$, and assuming a defect density $\rho_d = 10^{15}$ cm$^-3$. The multiphonon SRH self-energies were computed assuming carrier lifetimes $\tau_{n,0} = 9$ ns and $\tau_{p,0} = 46$ ns, and Fig. 3.7(a) shows

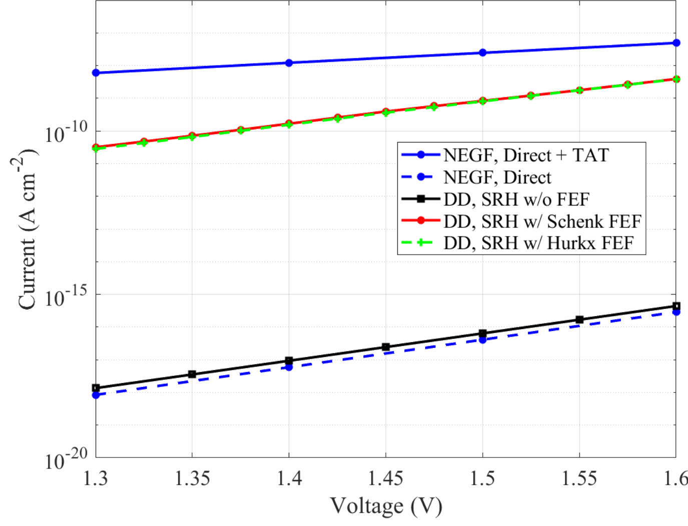


Figure 3-8: *I-V characteristic for the p-n junction computed with NEGF and drift-diffusion (DD). The LDOS tails generated by the high electric field at the junction enhance the recombination current, yielding an ideality factor* $\eta_{NEGF} \approx 5.5$*. If the self-energy is restricted to* Σ_{Direct}*, neglecting tail states contribution, the NEGF current coincides with the one obtained from DD* ($\eta_{DD} = 2$)*. DD simulations including Schenk and Hurkx field-enhancement factors (FEF) are also shown for comparison.*

the resulting spectral current density at $V_{\text{bias}} = 1.3$ V. The continuous stripes originating from the contacts, which represent current-carrying channels, become discontinuous at the defect position, where they switch band due to SRH recombination. Moreover, the observed stripes are separated in energy by $\hbar\omega_0$ steps, because multi-phonon recombination can only occur at integer multiples of the effective phonon energy with respect to the defect level. The current density shown in Fig. 3.7(b), which result from the energy integration of the spectral current density in Fig. 3.7(a), displays a digital switch between electron and hole currents at the defect position as it is expected because of SRH recombination, and the resulting total current is constant over the entire device, hence reflecting that the SRH self-energy satisfy the current conservation condition (2-61).

The resulting *I-V* characteristic for the p-n junction is shown in Fig. **3-8**, for bias voltages ranging between 1.3 and 1.6 V. These values where chosen because they correspond to the subthreshold region of the *I-V* curve, where most of the current comes from defect recombination. It can be observed that the NEGF current is greatly enhanced due to tunneling with respect to the semiclassical current obtained from the drift-diffusion simulation, where defect recombination is obtained from the usual SRH term in Eq. (3-36). The ideality factor

resulting from the NEGF simulation, which is computed through the formula

$$\eta = \frac{e}{k_B T}\left(\frac{\partial[\ln I]}{\partial V}\right)^{-1}, \tag{3-50}$$

gives $\eta_{\mathrm{NEGF}} \approx 5.5$, which reflects a deviation from the ideal behavior given by the drift-diffusion simulation and which is predicted by the Shockley theory $\eta_{\mathrm{DD}} = 2$. Such high values of the subthreshold ideality factor are considered a signature of the presence of defects, and are attributed to an enhancement of the defect recombination current by tunneling [13, 15]. In Fig. **3-8**, drift-diffusion simulations including Hurkx and Schenk field-enhancement factors are also shown for comparison, in which the tunneling masses used were the GaN electron and heavy hole effective masses[9]. As a final remark, it is interesting to observe from Fig. **3-8** that an additional NEGF simulation was performed using the only the direct component of both the defect and SRH self-energies, and the resulting current closely reproduces the result from the drift-diffusion simulation. This is in fact the expected result, since this component of the self-energy neglects the contribution from tail states assisting the tunneling process, and only takes into account direct thermal relaxation at the defect position, which is consistent with the semi-classical theory of defect recombination.

3.4 Simulation of an InGaN/GaN LED

The structure under consideration, shown in Fig. **3-9**, was taken from [12] (device A2). It is composed of an n-doped GaN buffer layer with $N_D = 1 \times 10^{20}$ cm^{-3}, a 4 nm GaN spacer (SP), a 3 nm $In_{0.17}Ga_{0.83}N$ quantum well (QW), a 4 nm GaN quantum barrier (QB), a 20 nm $Al_{0.13}Ga_{0.87}N$ electron-blocking layer (EBL) with $N_A = 5 \times 10^{19}$ cm^{-3} and a p-GaN contact region with the same doping concentration. Incomplete dopant ionization is assumed, with activation energies $\Delta E_D = 20$ meV and $\Delta E_A = 200$ meV for donors and acceptors, respectively. An 8-band $\mathbf{k} \cdot \mathbf{p}$ description of the electronic structure for wurtzite crystals is used, which includes the first conduction band, heavy-hole, light-hole and spin-orbit split-off bands (band parameters are given in Tab. **A-2** of App. A). The simulations are performed at room temperature, with the material parameters shown in Tab. **3-2**, and assuming a transversal area $A = 1$ mm^2. A linear interpolation between the values of binary compounds has been used for the ternary compound parameters, with the exception of E_g which includes a quadratic term of the form $x(1-x)b$, with the bowing parameters $b_{\mathrm{AlGaN}} = 0.7$ eV and $b_{\mathrm{InGaN}} = 3.0$ eV. Additionally, a conduction band discontinuity $\Delta E_c/\Delta E_g = 0.7$ has been assumed for all heterojunctions.

The effect of spontaneous and piezoelectric polarization charges at hetero-interfaces is important in III-nitride semiconductor with wurtzite crystal structure, as it can lead to strong

[9] We avoid writing the explicit expressions of the field-enhancement factors as it is not the scope of this work to study these expressions, but the interested reader is directed to Refs. [72, 73].

internal electric fields that may significantly increase carrier tunneling into defect states. Such effects were included in the solution of the Poisson's equation as detailed in Sec. 2.8.2. Defect parameters were chosen according to [75], where capacitance-voltage and steady-state photocapacitance measurements are used to characterize defects acting as non-radiative recombination centers in GaN-based LEDs, concluding that defects are located near midgap, with defect densities in the order of 10^{15} cm^{-3} that are maximum near In-rich layers with some spreading in the surrounding regions. For simplicity, we assume a uniform distribution $\rho_d = 1 \times 10^{15}$ cm^{-3} of defect located at midgap. The carrier lifetimes were chosen according to defect concentrations and cross-sections[10] reported in [12], yielding lifetimes in the order of 1-10 ns. In all our simulations the values $\tau_{n,0} = \tau_{p,0} = 1$ ns were assumed.

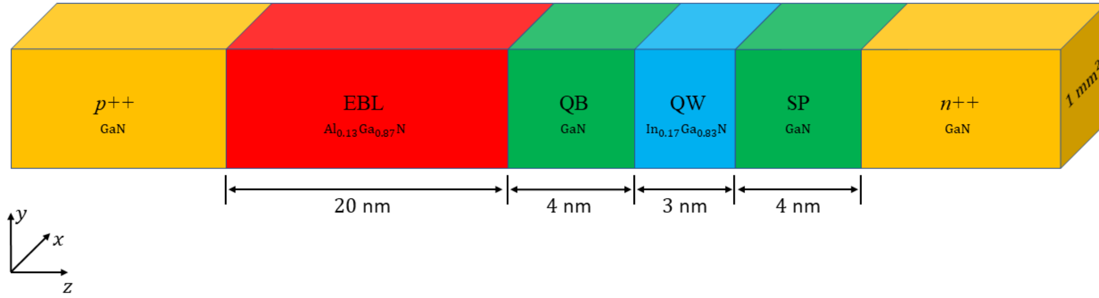


Figure 3-9: *Schematic of the SQW LED studied in this section (device A2 in Ref. [12]), with the corresponding physical dimensions of each layer.*

3.4.1 Numerical considerations

The very large heavy-hole effective mass present in III-nitride materials make it necessary to choose both a sufficiently small discretization spacing, and large $k_{\max}$ for the integration in the reciprocal space. The longitudinal space was discretized using an uniform mesh with spacing $\Delta z = 0.5$ nm, yielding good results but also imposing a limitation in the total length of the structure because of the high computational demand of the 8-band model. Due to this limitation, just a small portion of the contacts was included in the NEGF simulation region: 8 nm of the p-contact and 5 nm of the n-contact proved to be enough to achieve flat-band conditions at the boundaries within the range of bias voltages used in the simulations, yielding a total length of the device $L = 44$ nm in the growth direction.

Transverse momentum space was discretized using a uniform mesh with $k_{\max} = 3$ nm^{-1} and resolution $\Delta k = 0.0375$ nm^{-1}, which proved to be enough cover the spectral range of carrier and current density components with the given scattering parameters. Finally, the energy

[10]The capture cross-section $\sigma_{n(p)}$ is related to the average capture coefficient through the equation $\langle c_{n(p)} \rangle = v_{\mathrm{th},n(p)}\sigma_{n(p)}$, where $v_{\mathrm{th},n(p)}$ is the electron (hole) thermal velocity.

Table 3-2: Material parameters used in the simulation, taken from [76, 77, 78].

	Parameter	GaN	AlN	InN	Unit
Lattice parameter ($T = 300$ K)	a_L	3.189	3.112	3.545	Å
Lattice parameter ($T = 300$ K)	c_L	5.185	4.982	5.703	Å
Bandgap energy ($T = 300$ K)	E_g	3.44	6.16	0.69	eV
Electron effective mass (c-axis)	m_e^z	0.186	0.322	0.065	m_0
Electron effective mass (transversal)	m_e^t	0.209	0.329	0.068	m_0
Static dielectric constant	ϵ_s	9.5	8.5	15.0	ϵ_0
High frequency dielectric constant	ϵ_∞	5.35	4.6	6.7	ϵ_0
Polar optical phonon energy	$\hbar\omega_{\text{LO}}$	90	90	90	meV
Acoustic deformation potential	D_a	8	8	8	eV
Density	ρ	6.15	3.23	6.81	g/cm^3
Longitudinal sound velocity	u_l	6.56	6.56	6.56	10^5 cm/s
Screening length	q_0^{-1}	10	10	10	nm
Huang-Rhys factor	S	10	10	10	

range of the simulations was determined by adding a 0.5 eV range above and below the band edge extrema for each value of the bias voltage, with a uniform resolution of $\Delta E = 5$ meV that was enough to take into account all the sharp spectral features present in the QW region. The band structure and contact Fermi levels used as input in the NEGF simulation were determined through an initial drift-diffusion simulation of the structure.

3.4.2 Defect recombination in the subthreshold regime

We start our analysis of the SQW LED structure in the low forward bias region, where almost all the current flowing through the device can be attributed to defect recombination. An initial simulation was carried out at a forward bias of $V_{\text{bias}} = 2.0$ V, which is still well below the threshold voltage of the device but is high enough such that the computational burden can be reduced as a smaller energy range must be considered. Fig. **3-10** shows the local density of states (LDOS) computed with Eq. (2-84), for vanishing transverse momentum (*i.e.*, just the term $\mathbf{k} = 0$ is considered). The blue shades in the colour map, which are proportional to the magnitude of the LDOS at each point in the position-energy plane, are useful to qualitatively see the position of (quasi)bound states in the quantum well. The black contour lines of the LDOS indicate a decrease of two orders of magnitude in the tail states (not visible in a linear color map) away from the band edges. A cut of the ($\mathbf{k}$-integrated) LDOS at $z = 23$ nm is shown in the right panel of Fig. **3-10** (blue lines), where defect recombination is maximum, see below Fig. 3.11(a). Tail-state formation in the LDOS exhibit a complex interplay between field-induced tails, confined state resonances

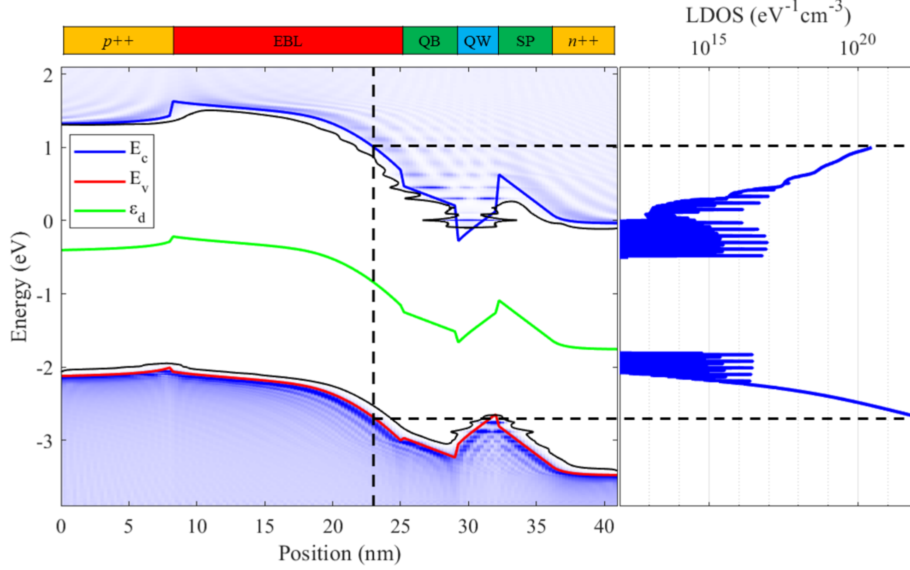


Figure 3-10: *Local density of states (left panel, color maps) of the single-quantum-well LED structure under investigation in weak forward bias $V_{bias} = 2.0$ V, showing a complex structure of bond and quasi-bond states. Contour lines (black curves) show the extension of these resonant states in the energy gap. The green line marks the energy level of the defect, which is assumed to lie at midgap (conduction and valence band edges are represented by blue and red lines, respectively) and uniformly distributed in the structure. Sharp phonon resonances related to multiphonon relaxation can be observed in the logarithmic plot of the LDOS evaluated at $z = 23$ nm (right panel), where defect recombination is maximum.*

extending into the energy gap (see the ladder shaped tail extending from the conduction band), and sharp phonon resonances related to multi-phonon relaxation.

Fig. 3.11(a) shows the net SRH recombination rate computed by inserting the SRH self-energy in Eq. (2-87), *i.e.*, through the formula

$$R_{\mathrm{SRH}}(z_i) = \frac{1}{A\Delta z}\int \frac{dE}{2\pi\hbar}\sum_{\mathbf{k}} \mathrm{Tr}\Big[\big\{\mathbf{\Sigma}^R_{\mathrm{SRH}}(\mathbf{k},E)\mathbf{G}^<(\mathbf{k},E) + \mathbf{\Sigma}^<_{\mathrm{SRH}}(\mathbf{k},E)\mathbf{G}^A(\mathbf{k},E) - \mathbf{G}^R(\mathbf{k},E)\mathbf{\Sigma}^<_{\mathrm{SRH}}(\mathbf{k},E) - \mathbf{G}^<(\mathbf{k},E)\mathbf{\Sigma}^A_{\mathrm{SRH}}(\mathbf{k},E)\big\}_{i,i}\Big] \quad (3\text{-}51)$$

with the energy integration restricted to either the conduction or valence band, and the trace is over band indices. Two peaks can be observed in the R_{SRH} profile, lying at $z_1 = 23$ nm and $z_2 = 32$ nm. The first peak corresponds to the cross-point between electron and hole densities (see Fig. 3.11(b) on the right-hand side), *i.e.*, where the product np is maximum, while the second peak rests at the position where the overlap between electron and hole quasibound states in the QW region is maximum[11]. Since most of the recombination occurs in the

[11] Notice from Fig 3.11(b) that electron and hole densities peak at different positions inside the QW region.

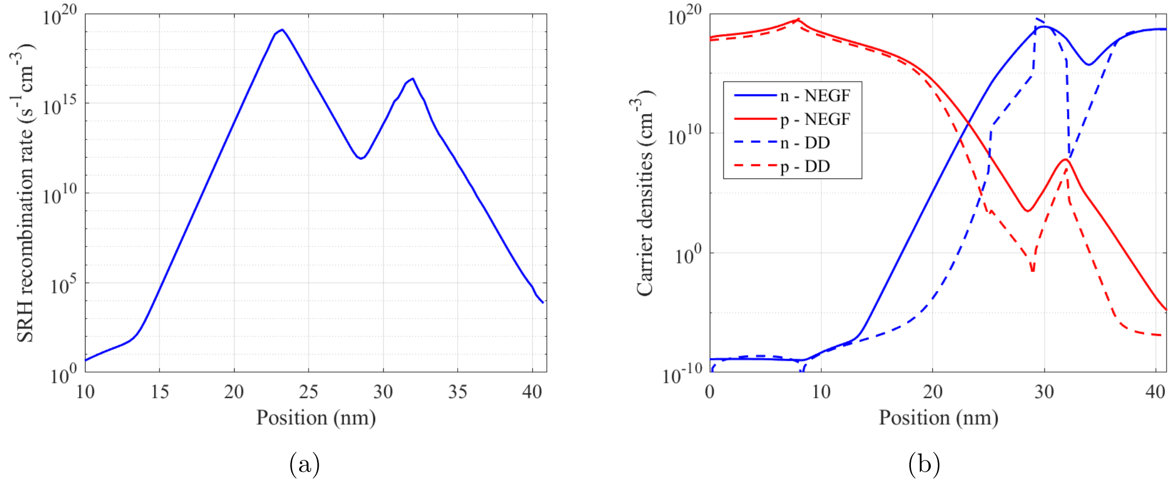


Figure 3-11: *(a) SRH recombination rate and (b) Carrier densities at a forward bias of $V_{bias} = 2.0$ V. The effect of tail states induced by the strong internal fields can be appreciated in the increment of the NEGF densities with respect to the semi-classical situation (drift-diffusion simulation, represented by dashed lines).*

high-electric-field regions, the contribution of carrier-phonon scattering is negligible with respect to the field-induced contribution; the Urbach tails due to deformation potential and polar optical scattering can be better appreciated in the flat-band regions (the quasineutral layers of the LED where the subgap states induced by the band bending are not present), as was already discussed in the previous section. Therefore, we may argue that the tunneling assisting SRH recombination is mainly a coherent process. However, as already observed in the case of BTBT [79], inelastic carrier-phonon scattering is responsible for the population of the confined states in the QW, which cannot be reached directly from the contacts, and thus indirectly contributes to TAT.

The SRH scattering current can be computed from the recombination rate through the equation

$$J_{\text{SRH}}(z, E) = e \int_{-\infty}^{z} dz' \, R_{\text{SRH}}(z', E), \tag{3-52}$$

where the spectral recombination rate $R_{\text{SRH}}(z, E)$ is obtained from Eq. (3-51), ignoring the integral over energy [80]. The resulting spectral current for different forward biases is shown in Fig. **3-12** (left panel) with their corresponding (energy-integrated) scattering current (right panel). Current-carrying channels (continuous horizontal stripes originating

This effect, known as the *quantum confined Stark effect*, is caused by the separation of electron and hole wavefunctions in the quasi bound state with increasing electric field.

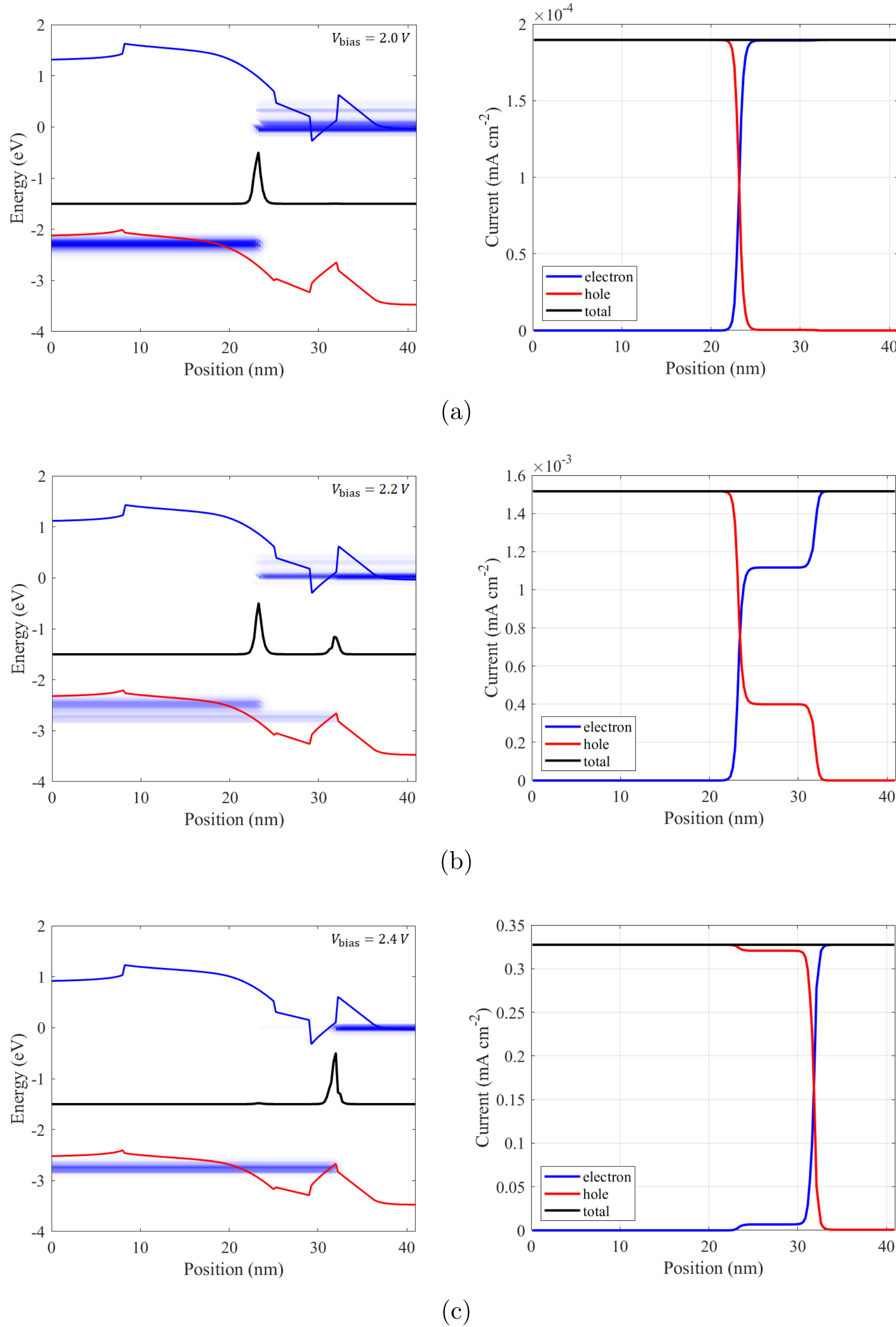


Figure 3-12: *Spectrally resolved current density (left panel) and the corresponding recombination current resulting from energy integration (right panel) computed with NEGF at a forward bias of (a) $V_{bias} = 2.0$ V, (b) $V_{bias} = 2.2$ V and (c) $V_{bias} = 2.4$ V. The continuous stripes in the spectral current, corresponding to current-carrying channels, switch bands where the SRH recombination rate is maximum. Black lines represent the corresponding energy-integrated recombination rates (in arbitrary units). The integrated currents on the right-hand side reflect how carriers gradually recombine when moving from the EBL-QB interface to the QW, while the total current is perfectly conserved.*

from the contacts), fade away when they reach the region where R_{SRH} (solid black line) is maximum. The position of the recombination window is the result of a complex interplay between tunneling probabilities and multi-phonon matrix elements. At a bias of 2 V, most of the recombination occurs near the EBL-QB heterojunction, as was pointed out in Fig. 3.11(a), which is also where the magnitude of the electric field is maximum (see below Fig. 3.13(a)). The electrons reach the EBL by resonant tunneling (the electron spectral current approximately flows at the ground and first excited state energies of the QW), while holes have to tunnel a shorter distance before reaching the virtual states that participate in the recombination process.

The second recombination peak at $z_2 = 32$ nm gradually increases with the applied bias, moving the recombination window towards the n-side of the junction. At 2.2 V, Fig. 3.12(b) shows the corresponding spectral current and (energy-integrated) recombination current, and it can be observed that the second peak acquire a magnitude comparable to the peak at $z_1 = 23$ nm (EBL-QB interface). As the quasi bound state energy in the valence band QW gets closer to the p-contact Fermi level, more holes are able to tunnel directly and recombine at the QW by multi-phonon relaxation (a second blue stripe can be observed in the valence band spectral current fading at z_2). This is also confirmed in the right-hand side panel, where the recombination current shows a flow of holes moving through z_1 and finally recombining at z_2. As the applied bias is further increased to 2.4 V, Fig. 3.12(c) shows that the recombination window is completely shifted to the second peak at z_2. This time both electrons and holes tunnel to the ground state in the QW and are subsequently captured by the defect state.

The shift in the recombination window from the first to the second peak can be traced back to the combination of two effects: first we have a decrease in the electric field in the EBL region with increasing bias voltage, as can be seen in Fig. 3.13(a). This behaviour, combined with the small increment in the electric field in the QW region, leads to a decrease in the tunnelling probability of electrons from QW into the EBL subgap states. On the other hand, a second cause can be identified by looking at the carrier densities as a function of the applied bias in Fig. 3.13(b), where the shift in the recombination window may be traced back to the slow increase in the hole population in the QW. The different bias dependencies of electron and hole population in the QW is indicative of inefficient hole injection, which has been reported as the origin of the inhomogeneous distribution of holes in multi-QW LEDs (see *e.g.*, Ref. [81]).

3.4.3 NEGF connection with the semiclassical SRH theory

In Sec. 3.1.1 it was demonstrated that, for a bulk semiconductor, the NEGF formulation of defect recombination is equivalent to the semi-classical SRH formulation in Eq. (3-36), only in quasi-equilibrium conditions, *i.e.*, when correlation and spectral functions are re-

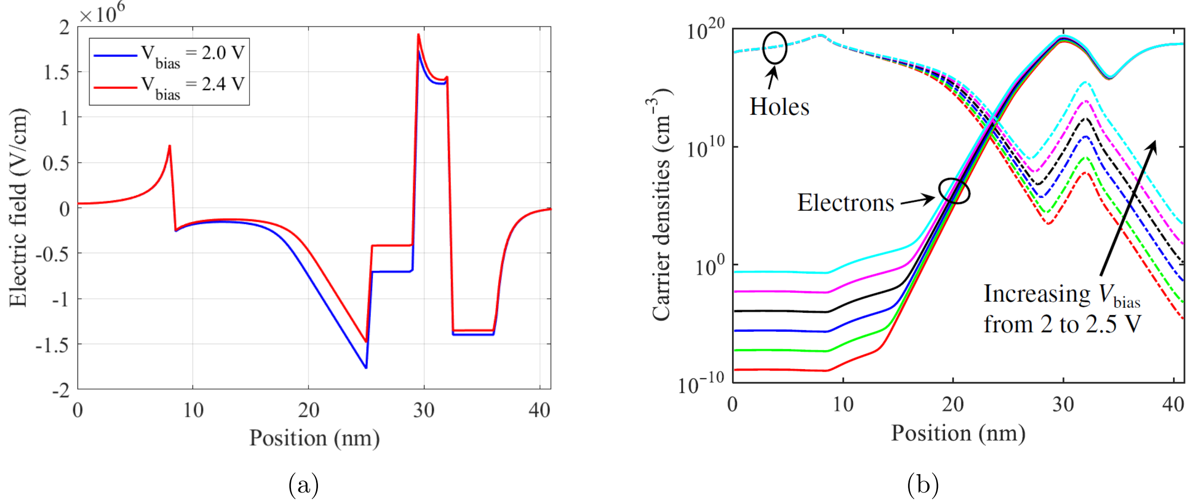


Figure 3-13: *(a) Electric field profile along the structure at a forward bias of* $V_{bias} = 2.0$ *V (blue line) and* $V_{bias} = 2.4$ *V (red line). (b) Electron and hole densities at different forward bias voltages.*

lated by the local fluctuation-dissipation theorem. Moreover, results from the test structure used in the previous section (GaN p-n junction) showed that defect recombination in NEGF restricted to the direct component of the SRH self-energy in Eqs. (3-48a)-(3-48b) yield a similar recombination current than the one obtained with the semi-classical SRH formula within a drift-diffusion simulation framework (see Fig. **3-8**). It is interesting to see whether the semi-classical SRH formula, with the correct quantum carrier densities, can still reproduce the field-enhanced SRH rate in a nanostructure like the one under consideration. Previously we found that the NEFG SRH rate peaks at two points $z_1 = 23$ nm and $z_2 = 32$ nm where the product np is maximum, which suggests that this may be the case.

The comparison between NEGF and the semi-classical SRH formulation is shown in Fig. **3-14**. The SRH rate, labelled as "*quantum-corrected SRH*" was computed through the formula

$$r_{\text{SRH}} = \frac{np - n_{\text{eq}}p_{\text{eq}}}{\tau_{p,0}(n+n_1) + \tau_{n,0}(p+p_1)}, \tag{3-53}$$

where n, p are the NEGF carrier densities, and the equilibrium NEGF densities n_{eq} and p_{eq} were used instead of the intrinsic carrier density n_i. The good agreement with NEGF calculations suggests that drift-diffusion solvers with a quantum-corrected LDOS obtained, *e.g.*, from the self-consistent solution of the Schrödinger equation [82] or the localization landscape theory [83], represent a viable approach to the analysis of TAT. Indeed, within these quantum-corrected DD schemes, TAT would be naturally included within the standard SRH theory, without the need of specific models for field enhancement factors, whose expres-

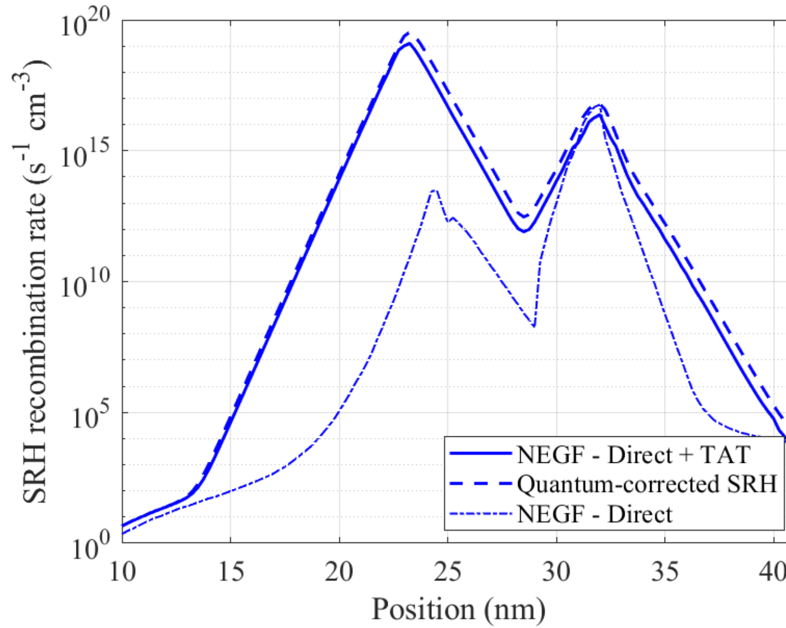


Figure 3-14: *Comparison of the SRH rate computed with different models at a forward bias of $V_{bias} = 2.0$ V. The SRH rate computed with NEGF (solid blue line) is well reproduced by the quantum-corrected SRH rate (dashed blue line) computed from Eq. (3-53) with the NEGF carrier densities. Neglecting tail state contribution, which corresponds to the classical limit (i.e., SRH without the inclusion of TAT), significantly underestimates the recombination rate (dotted blue line).*

sions tend to be overly complex, with many unknown quantities, in particular the tunneling masses, treated as fitting parameters [13].

3.4.4 Fitting of the I-V characteristic

The *I-V* characteristics of the LED in Fig. **3-9** was computed with NEGF, for an applied bias between 1-3 V. Fig. 3.15(a) shows a comparison of the results from the NEGF approach (blue line) and from experiments (blue circles). The experimental data was taken from Ref. [12] (device A2). The fitting of the experimental results is obtained by assuming that $\tau_{n,0} = \tau_{p,0} = 1$ ns, which is compatible with defect concentrations and cross sections reported in Ref. [12].

In the subthreshold regime, below the optical turn on, TAT is the most relevant contribution to carrier transport. Deviations from the ideal behaviour predicted by Shockley theory can be better appreciated by looking at the ideality factor, computed with Eq. (3-50), which is also shown in Fig. 3.15(a) (red line and circles). Ideality factors between one and two are

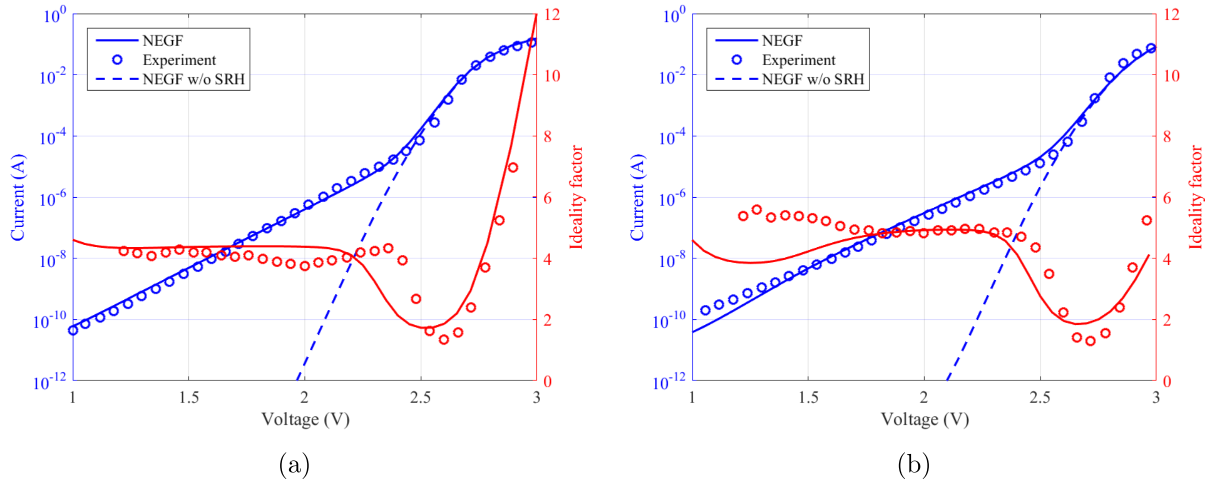


Figure 3-15: *(a) The I-V characteristics (left axis) and corresponding ideality factor (right axis), computed with the NEGF approach (solid lines) and from measurements in Ref. [12], device A2 (circles). A series resistance $R_s = 1\ \Omega$, accounting for nonideal contacts and buffer layers, is selected to match the slope of the experimental I-V curve at high voltages. (b) The simulation was repeated for the device A1 from Ref. [12], in which the In concentration in the QW is reduced to 15%.*

normally attributed to the competition between diffusion and the recombination processes, while ideality factors exceeding two are attributed to tunnelling. For bias voltages above 2.5 V, as the diffusion current (dashed blue line) becomes dominant, the ideality factor decreases, approaching one, before increasing again due to the resistive behaviour of the diode. A series resistance $R_s = 1\ \Omega$ was also added to account for the effect of nonindeal contacts and substrate.

The computations were also performed for the device A1 in [12], which feature the same physical dimensions as the previous example with the difference that the Indium content in the QW is 15%. The fitting was achieved with the same values for the carrier lifetimes, and a series resistance of $R_s = 0.3\ \Omega$. Once again, the NEGF results are able to accurately predict the I-V characteristics of the LED, and in particular, the high ideality factors in the subthreshold regime.

Chapter 4

NEGF modeling of type-II superlattice detectors

This section will be concerned with the modeling of type-II superlattice (SL) photodetectors using the NEGF approach. Transport simulations in broken-gap configurations calls for a full quantum kinetic model with a multiband description of the electronic dispersion relation, as it involves a complex interplay of several mechanisms ranging from coherent interband tunneling (through miniband formation) to sequential tunneling or Wannier-Stark hopping, depending on the built-in and/or externally applied field. Moreover, the nonlocality of carrier-photon interactions within the NEGF framework is essential for spatially indirect optical transitions in superlattice absorbers. In order to tackle the staggering computational effort coming from the self-consistent solution of a multiband implementation of NEGF in realistic devices featuring relatively large dimensions, a solution in mode-space will be used, where NEGF equations are transformed using a reduced set of problem-matched basis functions. Such basis set is obtained from the eigenfunctions of the noninteracting Hamiltonian of the nanostructure, which are then moved to an orthogonal basis set of functions by diagonalizing the position operator, similar to the procedure followed to find maximally localized Wannier functions in quantum cascade lasers [84, 85]. After establishing the main theoretical aspects of this mode-space approach, we will use NEGF to analyze an interband cascade infrared photodetector (ICIP) and demonstrate the accuracy of the proposed model. A semiclassical perspective to carrier transport in superlattices will be given by means of a mobility study. Finally, we will study an nBn absorber following a different approach, in which self-energies accounting for phonon scattering are computed within the Büttiker probes formalism, and generation-recombination mechanisms dealing with optical absorption and SRH are included directly in the NEGF continuity equation.

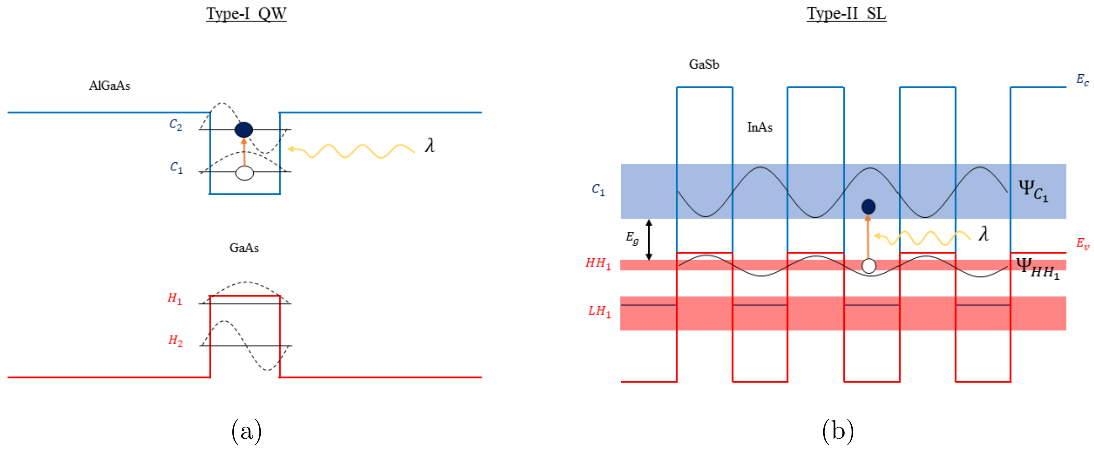


Figure 4-1: *Schematic illustration of infrared detectors based on (a) AlGaAs/GaAs type-I QWs and (b) GaSb/InAs type-II SLs. In the former case, IR absorption occurs through intersubband transitions between energy levels inside the conduction band QW (n-doped) or the valence band QW (p-doped). On the other hand, type-II SLs exploit miniband formation, where optical transitions occur between conduction and valence minibands, with the minimum transition energy defined by the effective bandgap determined by the difference between the first conduction miniband C_1 and the first heavy-hole miniband HH_1.*

4.1 Brief introduction to type-II superlattice infrared detectors

The birth of type-II SL technologies can be traced back to the work of Sai-Halasz *et al.* [86], published in 1977, in which the authors proposed and analysed a novel superlattice structure involving $In_{1-x}Ga_xAs/GaSb_{1-y}As_y$ heterostructures. The intrinsic properties displayed by this new technology led to further investigations into InAs/GaSb SL in order to overcome the existing limitations present in alternative technologies such as type-I SL detectors (made of GaAs/AlGaAs) and the most commonly used IR detectors based on mercury cadmium telluride (HgCdTe). While HgCdTe detectors offered excellent performance in the mid-wavelength (MWIR) and long-wavelength infrared (LWIR) regions, they had some drawbacks, including high manufacturing costs and complex fabrication processes. Additionally, HgCdTe detectors suffered from inherent material defects that limited their performance. The interested reader is directed to [87] for comprehensive review and in-depth analysis of HgCdTe IR detectors, and [88] is suggested for a more profound historical overview of the developments of type-II SL detectors.

Type-II SLs are characterized by large band offsets, comparable to the bandgap energies

of the bulk constituent materials[1], leading to a broken-gap alignment where electron and hole quantum wells are located in different material layers. As a result, electron and hole levels are separated in space, reducing the corresponding overlap of their wavefunctions when compared with type-I SL structures (see Fig. **4-1**). When material layers are sufficiently thin, adjacent confined levels overlap forming a miniband, *i.e.*, a quasi-continuous energy band in which carriers are characterized by Bloch states delocalized in space. Some of the main features of type-II SLs are summarized below [89]:

- The effective bandgap, defined by the difference between the bottom conduction miniband C_1 and the top valence miniband HH_1, can be varied by changing the thickness of the material layers. Superlattices made of GaInSb/InAs allow some degree of flexibility in the design of IR photodetectors for incident light in the SWIR, MWIR and LWIR range.

- Binary compounds in the 6.1 Å family are III-V materials, which means they are more robust and chemically stable when compared to II-VI semiconductors (HgCdTe). This is important from the point of view of manufacturability.

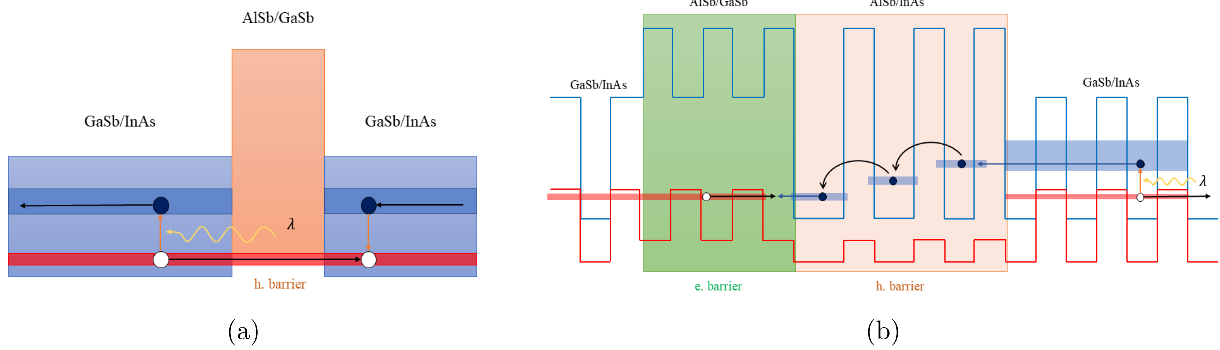
- Unlike the case of unstrained bulk materials, type-II SLs display a splitting between the heavy-hole band HH_1 and the light-hole band LH_1. This splitting can be increased (decreased) by applying compressive (tensile) strain to the superlattice.

- The HH_1 band is highly anisotropic and nearly dispersionless in the growth direction (*e.g.*, see Fig. **4-9**). This has profound implications in the transport properties of the SL, leading to extremely high heavy-hole effective masses and very low hole mobilities along the growth (transport) direction. As will be discussed latter in Sec. 4.4, one consequence of this is that hole transport is mainly a phonon-assisted process, where holes move through hopping between the confined levels inside the wells. This feature is also particularly unfavorable for the design of LWIR photodetectors, as thicker layers are necessary in order to reduce the effective bandgap of the SL.

- Electron effective masses tend to be heavier than that of the bulk materials. They are also weakly dependent on the SL design.

- Type-II SLs are more resistant to tunneling related leakage mechanisms (such as band-to-band tunneling and trap-assisted tunneling) when compared to IR detectors based on homojunction p-n diodes with narrow-gap semiconductors (with bandgap comparable to the SL effective gap). This is mainly due to the heavier electron effective mass in the transport direction [88].

[1] Usually GaSb, AlSb, InAs and their related compounds are used for IR applications, as they form a nearly lattice-matched family of semiconductors known as the 6.1 Å family [89].



(a) (b)

Figure 4-2: *(a) Schematic illustration of an nBn IR detector. In this example, an AlSb/GaSb SL barrier is sandwiched between an n-doped GaSb/InAs SL absorber. Minority holes generated by the light absorbed in the longer portion of the absorber recombine at the opposite side of the barrier, yielding a photocurrent through the device. (b) Schematic illustration of a single stage of a cascade IR detector. Multi-stage devices use both types of barriers in order to suppress leakage current. The hole barrier can also be engineered with wells of varying thicknesses in order to obtain a staircase-like disposition of the confined conduction band levels. Such disposition is used to facilitate the extraction of the photogenerated electrons from one stage into the following.*

- Auger recombination is expected to be greatly suppressed in comparison to HgCdTe detectors, mainly due to the spatially indirect transitions and the splitting between HH_1 and LH_1.

The huge variety of band offsets and bandgaps achieved by the semiconductor materials in the 6.1 Å family and their corresponding compounds allows for great flexibility and complexity in the different devices that can be constructed using this materials (including both type-I and type-II SLs structures). Among this configurations can be found nBn detectors (Fig. 4.2(a)), where nonradiative recombination mechanisms are highly suppressed using a SL barrier, or multi-stage cascade IR photodetectors/lasers[2] (Fig. 4.2(b)). The latter devices correspond to much more complex structures, including both electron and hole barriers engineered to suppress leakage currents and facilitate the extraction of the photogenerated e-h pairs, as illustrated in Fig. 4.2(b). In the following sections, our implementation of NEGF will be used to study both of this structures: an nBn detector will be studied in Sec. 4.5 and an interband cascade infrared photodetector will be briefly analysed in Sec. 4.3. Before of that, we will introduce our implementation of a mode-space approach for the NEGF formalism, which will greatly reduce the computational effort demanded by the numerical simulations.

[2]In antimonide-based IR lasers, the type-II SL used as active region in IR detectors, is replaced by a "W" type-II quantum well. This is because in this structure the overlap between electron and hole wavefunctions is increased, favouring optical generation. See Ref. [90].

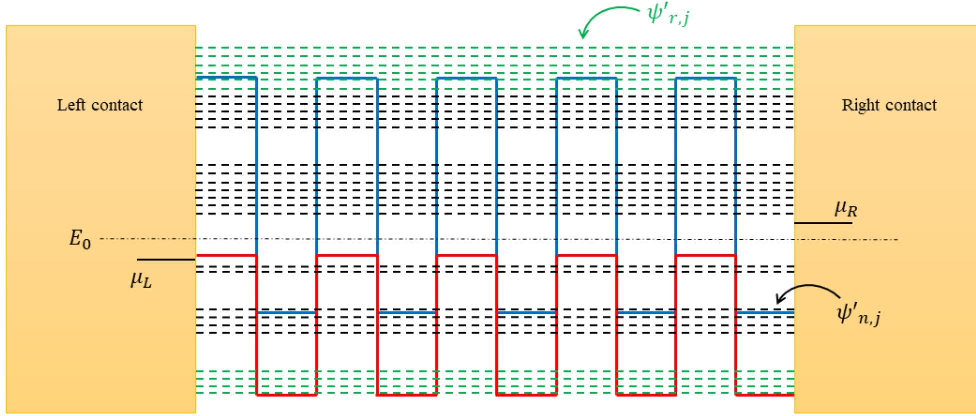


Figure 4-3: *Schematic illustration of the mode-space subdivision into near and remote basis functions for a SL structure. The eigenenergies of the noninteracting Hamiltonian $\boldsymbol{H}'_0$ (dashed lines), and their corresponding eigenfunctions, are classified as near $\psi'_{n,j}$ (black lines) and remote $\psi'_{r,j}$ (green lines), according to their proximity to some fixed value E_0 lying in the middle of the range considered.*

4.2 Mode-space analysis

We start by briefly describing the main theory concerning the mode-space approach, which we have already discussed in previous works, see Refs. [63, 91]. The main goal is to project the Green's function and self-energies into a lower-order suitable basis set, that is conveniently chosen to be the eigenfunctions of the noninteracting Hamiltonian of the nanostructure, which are then transformed into an orthogonal set by diagonalizing the position operator. A subset of the resulting eigenstates is then selected, spanning all the conductive channels of interest in the conduction and valence bands, *i.e.*, several $k_B T$ above and below the highest and lowest contact Fermi levels, respectively. The resulting subset of basis functions is referred as the *near* basis set, and the set of eigenfunctions outside the energy range of interest is referred as the *remote* basis set. In this representation, the number of basis functions required for a given accuracy should be much less than in the original space, reducing the computational effort significantly.

Starting from the original set $\{\phi_{im,\mathbf{k}}(\mathbf{r})\}$ (of dimension $N' = N \times N_b$) in the finite-element $\mathbf{k}\cdot\mathbf{p}$ basis, the eigenvalues of the noninteracting Hamiltonian $\mathbf{H}_0 = \mathbf{H}(\mathbf{k} = 0)$, from Eq. (2-106), are computed by solving the generalized eigenvalue problem

$$\mathbf{H}_0 \psi_j = E_j \mathbf{S} \psi_j, \tag{4-1}$$

with $\mathbf{S}$ the overlap matrix in Eq. (2-71), and $j = 1, 2, \ldots, N'$, labelling the the eigenvalues E_j. Löwdin symmetric orthogonalization is used to transform Eq. (4-1) into a standard

eigenvalue problem with symmetrically orthogonalized functions, closest in the least square sense to the original nonorthogonal functions, thus eliminating the need of different representations of the GFs (*i.e.*, covariant and contravariant, see Sec. 2.4)

$$\mathbf{H}_0'\psi_j' = E_j\psi_j', \tag{4-2}$$

with $\mathbf{H}_0' = \mathbf{S}^{-1/2}\mathbf{H}_0\mathbf{S}^{-1/2}$ and $\psi_j' = \mathbf{S}^{1/2}\psi_j$. The set of solutions of Eq. (4-2) are subsequently classified as near ($\psi_{n,j}'$) or remote ($\psi_{r,j}'$) as follows

$$\{\psi_{n,j}'\} = \{\psi_j' : |E_j - E_0| \leq \Delta\} \tag{4-3a}$$

$$\{\psi_{r,j}'\} = \{\psi_j' : |E_j - E_0| > \Delta\}, \tag{4-3b}$$

according to the proximity (Δ) of their respective eigenvalue E_j to some value E_0 lying in the middle of the energy range considered. The parameters E_0 and Δ are chosen such that the selected eigenenergies in the near basis cover all the relevant conductive channels of the system, which should include at least several k_BT units above and below the contact Fermi levels (see Fig. **4-3**). Analogous to the procedure followed to define maximally localized Wannier functions [84], we proceed to diagonalize the position operator in the near and remote representations to obtain maximally localized basis functions

$$\tilde{\mathbf{Z}}_\alpha\psi_{\alpha,j}^{\mathrm{ML}} = z_{\alpha,j}\psi_{\alpha,j}^{\mathrm{ML}}, \tag{4-4}$$

with $\tilde{\mathbf{Z}}_\alpha = \mathbf{T}_\alpha^\dagger\mathbf{S}^{-1/2}\mathbf{Z}\mathbf{S}^{-1/2}\mathbf{T}_\alpha$, where $\mathbf{Z}$ is the position operator in the finite-element basis, and $\mathbf{T}_\alpha = \{\psi_{\alpha,j}'\}_{j=1}^{N_\alpha}$ is the transformation matrix whose columns are the near ($\alpha = n$) or remote ($\alpha = r$) eigenstates. By defining $\mathbf{T} = \{\mathbf{T}_n, \mathbf{T}_r\}$ and $\mathbf{P} = \{\mathbf{P}_n, \mathbf{P}_r\}$, with $\mathbf{P}_\alpha = \{\psi_{\alpha,j}^{\mathrm{ML}}\}_{j=1}^{N_\alpha}$ the transformation matrix whose columns are the maximally localized basis functions, the Dyson and Keldysh equations read in the complete (near plus remote) space[3]

$$\left[E\mathbb{1} - \tilde{\mathbf{H}}(\mathbf{k}) - \tilde{\boldsymbol{\Sigma}}^R(\mathbf{k},E)\right]\tilde{\mathbf{G}}^R(\mathbf{k},E) = \mathbb{1} \tag{4-5}$$

$$\tilde{\mathbf{G}}^<(\mathbf{k},E) = \tilde{\mathbf{G}}^R(\mathbf{k},E)\tilde{\boldsymbol{\Sigma}}^<(\mathbf{k},E)\tilde{\mathbf{G}}^A(\mathbf{k},E), \tag{4-6}$$

where the GFs and self-energies are given by

$$\tilde{\mathbf{G}}^{R,<} = \mathbf{P}^\dagger\mathbf{T}^\dagger\mathbf{S}^{1/2}\mathbf{G}^{R,<}\mathbf{S}^{1/2}\mathbf{T}\mathbf{P} \tag{4-7a}$$

$$\tilde{\boldsymbol{\Sigma}}^{R,<} = \mathbf{P}^\dagger\mathbf{T}^\dagger\mathbf{S}^{-1/2}\boldsymbol{\Sigma}^{R,<}\mathbf{S}^{-1/2}\mathbf{T}\mathbf{P}, \tag{4-7b}$$

and the noninteracting Hamiltonian reads $\tilde{\mathbf{H}} = \mathbf{P}^\dagger\mathbf{T}^\dagger\mathbf{S}^{-1/2}\mathbf{H}\mathbf{S}^{-1/2}\mathbf{T}\mathbf{P}$. Denoting $\tilde{\mathbf{A}} = E\mathbb{1} - \tilde{\mathbf{H}} - \tilde{\boldsymbol{\Sigma}}^R$, the Dyson's equation (4-5) may be written in block matrix notation as

$$\begin{bmatrix} \tilde{\mathbf{A}}_{nn} & \tilde{\mathbf{A}}_{nr} \\ \tilde{\mathbf{A}}_{rn} & \tilde{\mathbf{A}}_{rr} \end{bmatrix}\begin{bmatrix} \tilde{\mathbf{G}}_{nn}^R & \tilde{\mathbf{G}}_{nr}^R \\ \tilde{\mathbf{G}}_{rn}^R & \tilde{\mathbf{G}}_{rr}^R \end{bmatrix} = \begin{bmatrix} \mathbb{1}_n & \mathbf{0} \\ \mathbf{0} & \mathbb{1}_r \end{bmatrix}, \tag{4-8}$$

[3]Notice the tilde symbol used here for GFs ($\tilde{G}$) and self-energies ($\tilde{\Sigma}$) refers to the operators expressed in the basis of maximally localized basis functions $\{\psi_{\alpha,j}^{\mathrm{ML}}\}$, not to be confounded with the covariant representation notation.

where each block $\tilde{\mathbf{G}}^R_{\alpha\beta} = \mathbf{P}^\dagger_\alpha \mathbf{T}^\dagger_\alpha \mathbf{S}^{1/2} \mathbf{G}^R \mathbf{S}^{1/2} \mathbf{T}_\beta \mathbf{P}_\beta$, with $\alpha, \beta \in \{n, r\}$, is a projection of the retarded GF into the near and remote basis using the respective transformation matrices (similarly for $\tilde{\mathbf{A}}_{\alpha\beta} = \mathbf{P}^\dagger_\alpha \mathbf{T}^\dagger_\alpha \mathbf{S}^{-1/2} \mathbf{A} \mathbf{S}^{-1/2} \mathbf{T}_\beta \mathbf{P}_\beta$). We are interested in the near basis representation of the GF, $\tilde{\mathbf{G}}^R_{nn}$, which contains the main spectral contribution of the scattering mechanisms to the system dynamics, so we may include the influence of remote states through a folding procedure, similar to the process followed in the Sec. 2.6 for the boundary self-energies. The final form of the kinetic equations in the near basis representation then reads

$$\left[E\mathbb{1} - \tilde{\mathbf{H}}_{nn}(\mathbf{k}) - \tilde{\mathbf{\Sigma}}^R_{nn}(\mathbf{k}, E) - \hat{\mathbf{\Sigma}}^R_{nn}(\mathbf{k}, E)\right] \tilde{\mathbf{G}}^R_{nn}(\mathbf{k}, E) = \mathbb{1}_n \tag{4-9a}$$

$$\tilde{\mathbf{G}}^<_{nn}(\mathbf{k}, E) = \tilde{\mathbf{G}}^R_{nn}(\mathbf{k}, E) \left[\tilde{\mathbf{\Sigma}}^<_{nn}(\mathbf{k}, E) + \hat{\mathbf{\Sigma}}^<_{nn}(\mathbf{k}, E)\right] \tilde{\mathbf{G}}^A_{nn}(\mathbf{k}, E), \tag{4-9b}$$

in which we have defined an additional self-energy $\hat{\mathbf{\Sigma}}^{R,<}$ representing the contribution of the remote modes to the coherent solution, given by (the dependence on transverse wavevector and energy has been dropped for clarity)

$$\hat{\mathbf{\Sigma}}^R_{nn} = \tilde{\mathbf{A}}_{nr} \tilde{\mathbf{g}}^R_{rr} \tilde{\mathbf{A}}_{rn} \tag{4-10a}$$

$$\hat{\mathbf{\Sigma}}^<_{nn} = -\tilde{\mathbf{\Sigma}}^{<B}_{nr} \tilde{\mathbf{g}}^A_{rr} \tilde{\mathbf{A}}^*_{rn} - \tilde{\mathbf{A}}_{nr} \tilde{\mathbf{g}}^R_{rr} \tilde{\mathbf{\Sigma}}^{<B}_{rn} + \tilde{\mathbf{A}}_{nr} \tilde{\mathbf{g}}^<_{rr} \tilde{\mathbf{A}}^*_{rn}. \tag{4-10b}$$

Here $\tilde{\mathbf{g}}^R_{rr} = \tilde{A}^{-1}_{rr}$, $\tilde{\mathbf{g}}^<_{rr} = \tilde{\mathbf{g}}^R_{rr} \tilde{\mathbf{\Sigma}}^{<B}_{rr} \tilde{\mathbf{g}}^A_{rr}$ and $\tilde{\mathbf{\Sigma}}^{<B}_{\alpha\beta}$ is the projection of the lesser boundary self-energy into the near and remote basis. From this point forward, the NEGF algorithm proceeds as usual, with the scattering self-energies computed directly in mode space, *i.e.*, replacing the node coordinates in the (electron-phonon) self-energies (2-152) and (2-153) with the eigenpositions $z_{n,j}$ from Eq. (4-4) [92]. Once self-consistency is achieved between GFs and self-energies, the scattering self-energies $\tilde{\mathbf{\Sigma}}^{R,<}_{nn}$ are transformed back to real-space, and the Dyson and Keldysh equations are solved once more in real-space to obtain the relevant one-particle properties. The effectiveness of the mode-space approach obviously depends on the number of basis functions needed to accurately represent the self-energies and GFs. In the next section, we will perform an accuracy assessment of the present approach, in which we will study a relatively large structure including only few modes in the near basis set. As we will demonstrate, the included self-energy term is able to regain current conservation as compared with the simple projection scheme, in which remote basis contribution is neglected.

4.3 Accuracy assessment of the mode-space approach on an ICIP device

We previously discussed the accuracy of the mode-space approach in Sec. 3.5.1 of [63] for an InAs/GaSb type-II SL using a 2-band $\mathbf{k}\cdot\mathbf{p}$ model, proving the equivalence with the standard

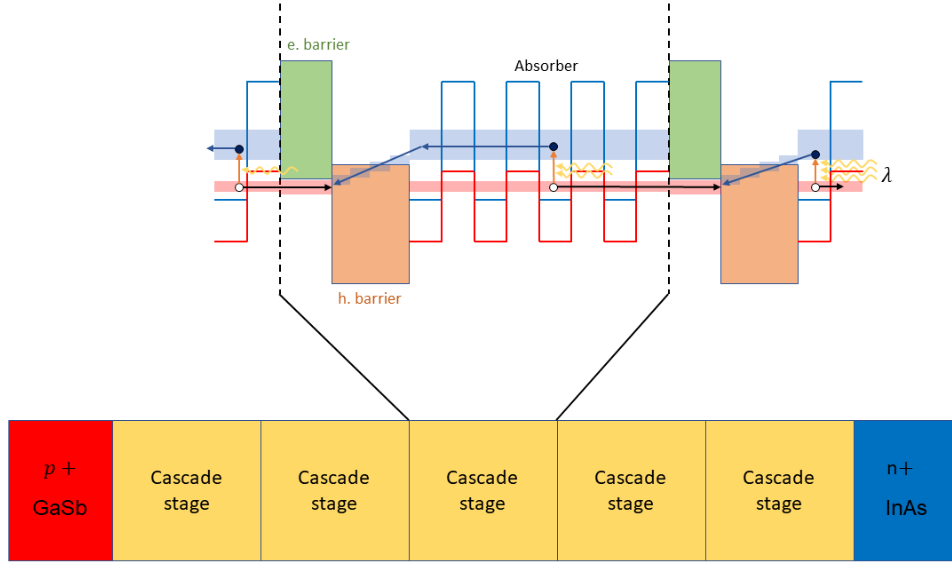


Figure 4-4: *Schematic of an ICIP device. It is composed of several cascade stages, containing type-II SL absorbers sandwiched between quantum-engineered barriers to avoid leakage current into subsequent stages. In this work we focus on a single cascade stage, composed of a GaSb/AlSb e. barrier, an InAs/AlSb h. barrier and an InAs/GaSb IR absorber.*

real-space NEGF algorithm while reducing the computation time and memory requirements of the simulation and also ensuring the conservation of the total current[4]. We proceed further here by testing the model in a complex type-II SL-based structure, encompassing all the possible carrier transport mechanisms, including interband tunneling, thus representing a good candidate for the accuracy assessment of the proposed mode-space approach.

Interband cascade infrared photodetectors achieve high-temperature and high-speed operation by employing a discrete absorber architecture, each single absorber interposed between quantum-engineered electron and hole barriers to form a series of interband cascade stages, as it is schematically depicted in Fig. **4-4**. At high temperatures, the diffusion length is typically shorter than the absorption depth of the infrared radiation. While the total thickness of the cascade can be comparable or even longer than the diffusion length, the photogenerated electrons travel only over one stage, which is significantly shorter than the diffusion length, before they recombine with the holes in the next stage. We will focus here on the structure of a single cascade stage (see Fig. **4-4**), including engineered barriers, designed to suppress intraband-tunneling current between different stages (*i.e.*, leakage current), and a p-doped ($N_A = 5 \times 10^{18}$ cm^{-3}) type-II InAs/GaSb SL absorber [93]. The *electron bar-*

[4]Mode-space approaches based on simple projection schemes, *e.g.*, the low rank approximation (LRA) method [92], improve NEGF computations in terms of time and memory, but violate current conservation due to the incomplete representation of the boundary self-energies.

rier consists of a p-doped 3 nm GaSb layer followed by two AlSb/GaSb QWs (16/52 Å). On the other hand, the *hole barrier*, acting also as a *electron relaxation ladder*, is made of a 9 period InAs/AlSb SL with 20 Å AlSb layers and InAs layers of varying dimensions: 88/80/72/64/56/48/40/36/32 Å, respectively. Finally, the *IR absorber* is composed of an 11 period InAs/GaSb SL (24/28 Å). The structure was simulated using an 8-band $\mathbf{k}\cdot\mathbf{p}$ model, including electron interaction with acoustic and polar optical phonons, and electron-photon interaction through an external source of monochromatic light that is assumed to be incident on the device. The material parameters used for the simulation are shown in Tab. **4-1**, while the Luttiunger parameters for the device Hamiltonian are shown in Tab. **A-1** from App. A.

Table 4-1: Material parameters used in the simulation, taken from [35, 94].

	Parameter	GaSb	InAs	InSb	AlSb	Unit
Lattice parameter	a_L	6.0854	6.0522	6.4717	6.1297	Å
Bandgap energy	E_g	0.8	0.407	0.227	2.374	eV
Electron effective mass	m_e	0.0412	0.0224	0.0135	0.13	m_0
Static dielectric constant	ϵ_s	15.7	15.15	16.8	12.04	ϵ_0
High frequency dielectric constant	ϵ_∞	14.4	12.3	15.7	10.24	ϵ_0
Polar optical phonon energy	$\hbar\omega_{\mathrm{LO}}$	30	30	30	30	meV
Acoustic deformation potential	D_a	8	8	8	8	eV
Density	ρ	5.61	5.68	5.77	4.29	g/cm^3
Longitudinal sound velocity	u_l	3.97	3.97	3.97	3.97	10^5 cm/s
Screening length	q_0^{-1}	10	10	10	10	nm
Valence band offset energy	ΔE_v	0	-0.56	0.03	-0.38	eV

The discretization parameters chosen for the simulation were $\Delta z = 0.5$ nm, $\Delta E = 5$ meV, $E_{\min} = -1$ eV, $E_{\max} = 1$ eV, $k_{\max} = 2.0$ nm^{-1} and $N_k = 100$. The length of the structure $L = 145$ nm, together with the discretization spacing chosen, translate into a total dimension of $N' = N \times N_b = 1160$ basis functions in our finite-element representation of the GF operators. Mode-space calculations will be performed using a total of $N_n = 250$ maximally localized basis functions, *i.e.*, the matrix rank will be reduced to approximately 22% of the original space. Maximally localized functions in the near basis set are selected according to their corresponding energy eigenvalues from Eq. (4-2), selecting those closest to $E_0 = 0$. The device was simulated at equilibrium conditions ($V_{\text{bias}} = 0$ V) and at room temperature.

In Fig. **4-5** we show the local density of states of the ICIP under consideration[5]. The hole and electron barriers also serve as electron-relaxation and interband-tunneling regions, respectively. The electron-relaxation region is designed to facilitate the transport of photo-

[5]The energy reference is set at the Fermi energy of the contacts, *i.e.*, $\mu_L = \mu_R = 0$ eV.

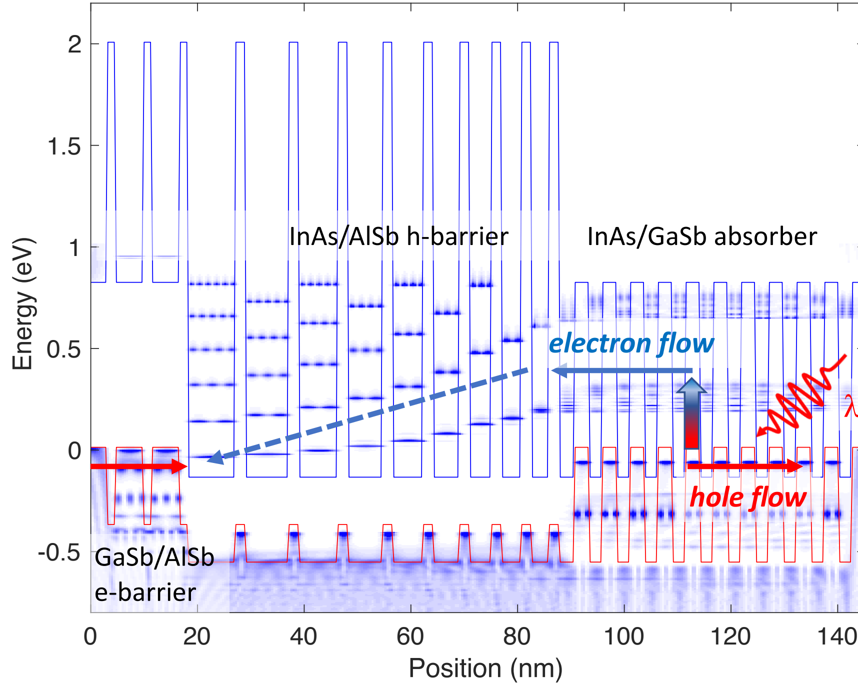


Figure 4-5: *Local density of states of one stage of a midwavelength IR interband cascade photodetector shown for* $\boldsymbol{k} = 0$ *(color map). The arrows indicate the direction of the electron and hole current flows: the electron-hole pairs photogenerated in the superlattice absorber diffuse along the respective minibands, the electrons towards the hole barrier, where they relax by polar optical transitions down a Wannier-Stark ladder, until they reach the electron barrier, whereupon they tunnel into the valence miniband of the adjacent stage.*

generated carriers from the conduction miniband of one stage to the valence miniband of the next[6]. For simplicity, the structure is considered in flat-band conditions, in order to avoid the difficulties arising from the ambiguous classification of electron and hole states when conduction and valence bands strongly intermix. Such classification is necessary for the correct definition of the electron and hole densities that will enter in the Poisson equation during the outer-loop iterations. Nevertheless, the inclusion of space-charge effects does not affect the assessment of the mode-space approach[7], that is our main interest during this first stage of the analysis. Moreover, ICIP devices are usually operated close to flat-band conditions [93], so this approximations is already able to give a good insight into the device behaviour at realistic conditions. Space-charge effects will be included later in this chapter, when we

[6]Operating under zero bias, ICIPs create one electron from a number of photons equal to the number of the constituent stages, which is exactly the opposite of quantum cascade lasers, and the photosignal is determined by the stage with the smallest photocurrent, usually the last one from the illuminated contact.

[7]A charge self-consistent electronic-structure model for type-II broken-gap SLs based on a multiband $\mathbf{k}\cdot\mathbf{p}$ envelope-function framework can be found in Ref. [95].

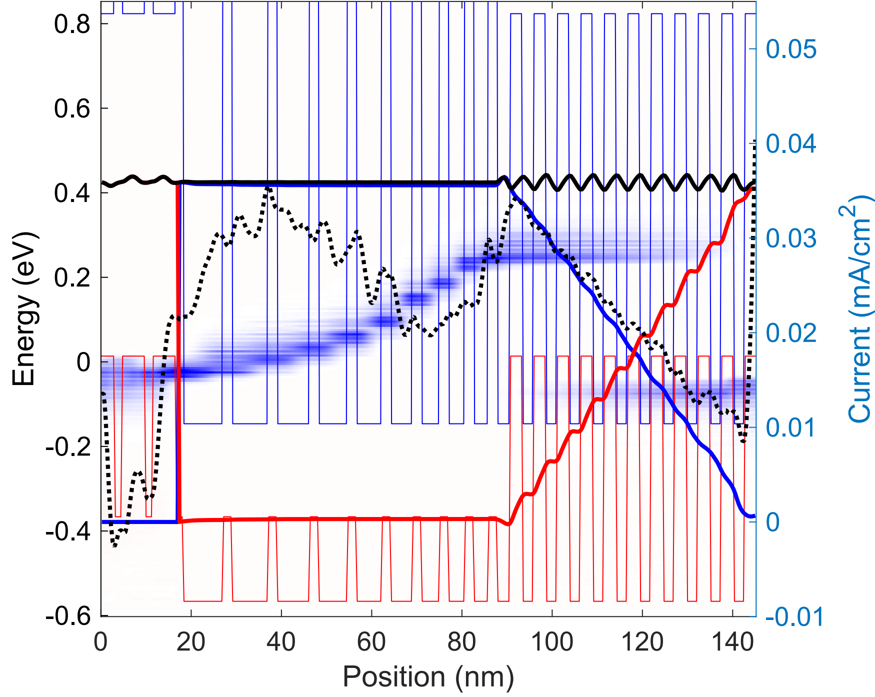


Figure 4-6: *Spatially resolved photocurrent spectrum under a monochromatic illumination of 1 W/m^2 with $E_\gamma = 0.3$ eV (color map). The corresponding energy-integrated current (black solid line, right axis) is approximately conserved, save for small oscillations due to the incompleteness of the basis set. Blue and red lines indicate electron and hole contributions, respectively. A position dependent energy threshold is defined according to the local density of states to separate electron and hole contributions (solid blue and red lines, respectively). Current conservation is clearly violated if the contribution of the remote modes is neglected (black dotted line).*

study transport across a barrier in an nBn superlattice structure in Sec. 4.5.

In Fig. **4-6**(a) we show the spectrally resolved photocurrent density in the single stage ICIP illuminated by a monochromatic light with intensity $I_\gamma = 1$ W/m^2 and photon energy $E_\gamma = 0.3$ eV. The continuous stripes in the current spectrum in the absorber region are indicative of miniband transport of the photogenerated electrons mediated by the extended states of the superlattice, while the staircase behavior in the relaxation layer is the signature of sequential tunneling assisted thermally by polar optical phonon relaxation. The interband tunneling through the type-II broken gap SL and finally the recombination with the holes of the next stage is also clearly visible. Upon convergence, the total energy-integrated current (black solid line) is approximately conserved over the whole device region, which is a very stringent test for any NEGF model. Residual fluctuations in the coherent current shown in Fig. **4-6**(a) (black solid line) may be traced back to the incompleteness of the basis when

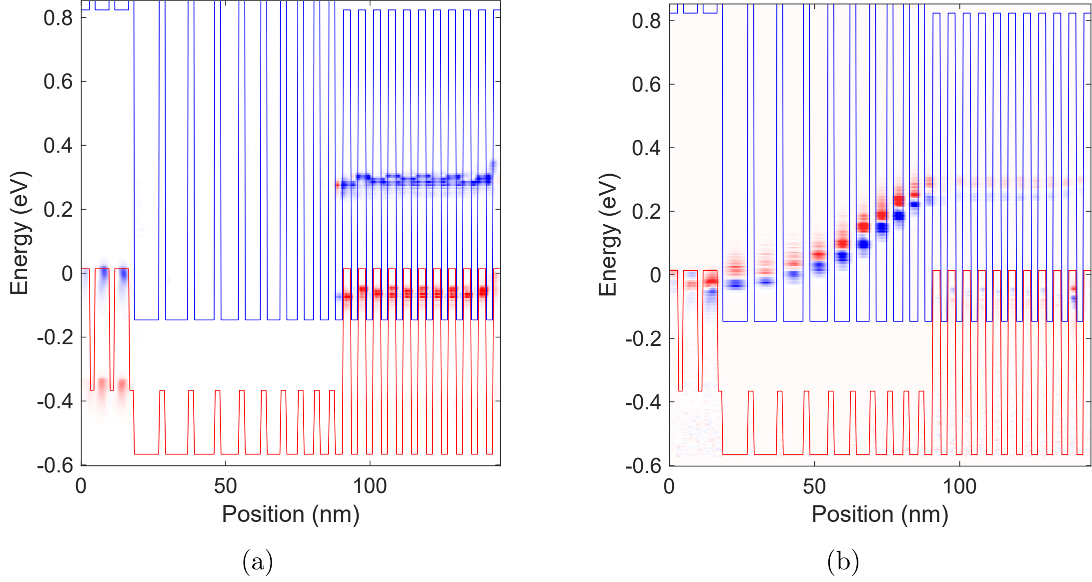


Figure 4-7: *(a) Spectrally resolved radiative generation rate (color map) and (b) spectral scattering current due to coupling to polar optical phonons, showing the relaxation of the carriers. Positive (negative) values, i.e., in (out) scattering of electrons, have a gradient from white to blue (red).*

GFs are transformed to the near basis set in the mode-space. Neglecting the contribution of the remote modes, *i.e.*, when the additional self-energies $\hat{\mathbf{\Sigma}}$ in Eqs. (4-10a)-(4-10b) are not included, results in severe violation of current conservation (black dotted line). Since the self-energies are additive, their contribution to the total current can be separated by inserting the self-energy of each scattering mechanism into Eq. (2-87) for the divergence of the current density (this is an approximation, of course, because GFs determine self-energies and vice versa) [80]. The radiative generation spectrum in Fig. 4.7(a), computed from Eq. (2-87) by inserting the electron-photon self-energy, shows the photogeneration of carriers in the superlattice absorber, and, unexpectedly, free-carrier absorption in the valence band of the electron barrier. The generation current

$$J_{\text{gen}}(z) = e \int_{-\infty}^{z} dz' \, G_{\text{opt}}(z'), \tag{4-11}$$

where the local generation rate $G_{\text{opt}}(z)$ is obtained from the energy integration of the radiative generation spectrum over a single miniband (either conduction or valence) [96], is found to give the same total current as in Fig. **4-6**, as it is expected in the case of complete carrier extraction in the absence of any recombination mechanism (*e.g.*, SRH, Auger). A comprehensive simulation of type-II SLs requires the inclusion of defect-mediated nonradiative recombination processes, which determine the extraction efficiency, but it was avoided

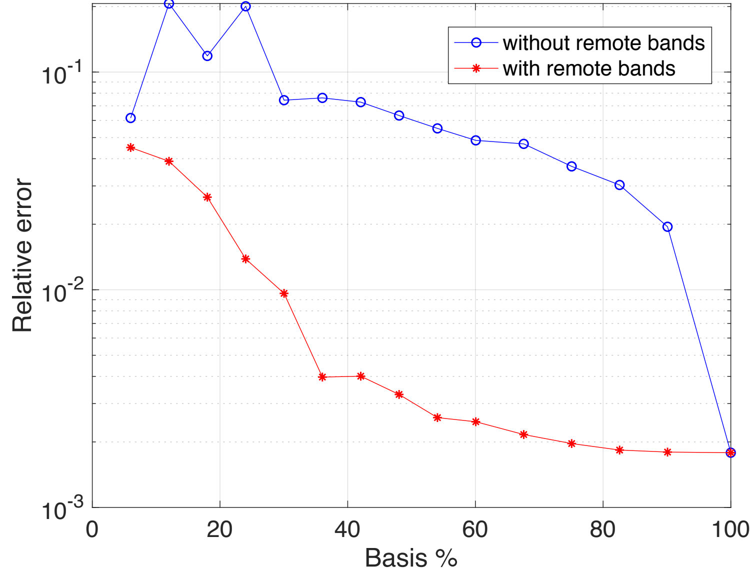


Figure 4-8: *Standard deviation of the photocurrent fluctuations normalized with respect to the mean spatial average, as a function of* $(N_n/N')^3$*, with and without the contribution of the remote modes.*

in the present study due to the enormous computational effort required by ICIP structure. Fig. 4.7(b) shows the spectral scattering current density (color map) associated to the coupling with polar optical phonons (similar to the previous case, Eq. (2-87) was used here with the polar optical phonon self-energy). The phonon-assisted cascade of the (minority) electrons from the conduction miniband of the superlattice absorber across the graded superlattice (relaxation region) is clearly visible.

Encoding the contribution of the remote modes in the self-energy from Eqs. (4-10a)-(4-10b) improves significantly the convergence with respect to conventional projection-based approaches. Because of memory limitations of our parallel OpenMP implementation, a convergence analysis with respect to the partition of the modes in near and remote sets is feasible only for the InAs/GaSb absorber, which can be observed in Fig. **4-8**. Convergence is better assessed on current densities, which are more sensitive to model-order reduction than other single-particle properties such as carrier densities. The speed-up in computation time is close to $(N_n/N')^3$ (*i.e.*, number of basis functions in the near set divided by the total number of basis functions), as Eqs. (4-10a)-(4-10b) are evaluated only once at the beginning of the self-consistent cycle (a computational complexity analysis of mode-space approaches is presented in Ref. [92]).

4.4 Connection to semiclassical approaches: mobility study in a LWIR type-II SL

We now turn our attention to the interpretation of the NEGF results from a semiclassical perspective. As a genuine quantum transport model, NEGF provides a comprehensive tool for the analysis of type-II SL photodetectors, including transport and generation-recombination processes through the corresponding energy- and momentum-dependent scattering self-energies. Nevertheless, it is interesting and useful to explore the connection between NEGF models and semiclassical approaches. Macroscopic quantities such as carrier mobilities and lifetimes are admittedly not germane to the NEGF formalism, but they represent critical ingredients needed in quantum-corrected drift-diffusion (DD) models [97]. A notable example of quantum-corrected DD approaches is the localization landscape theory, in which coherent processes such as localization and tunneling effects are included by means of an effective potential derived from a (non-autonomous) Schrödinger-like equation, while dissipative processes have to be described by appropriate mobilities [83, 98]. A mobility study could also be the starting point to develop fractional DD approaches, which may offer a unified description of carrier transport in disordered SLs, accounting for memory effects associated to trapping mechanisms, along the lines of recent investigations on anomalous (non-Fickian) transport in disordered organic semiconductors [99].

Different techniques were proposed to extract the mobility from quantum transport simulations. A dispersive, time-dependent diffusion coefficient was obtained from Monte Carlo simulations of AlAs/GaAs disordered SLs [101]. An "apparent" position-dependent mobility was proposed in a study of diffusive transport in fully depleted silicon-on-insulator transistors, bridging the gap between a Scharfetter-Gummel DD and NEGF models [102]. An efficient method to extract an effective mobility from NEGF and Kubo-Greenwood calculations, while minimizing contact resistance contamination and channel length misestimates, was discussed in the context of thin silicon films [103]. In the context of type-II SLs, electron and hole mobilities may be computed from a "drift experiment", by applying a small electric field F and measuring the resulting (average) carrier velocities

$$v_{n(p)} = \left\langle \frac{J_{n(p)}}{\rho_{n(p)}} \right\rangle, \tag{4-12}$$

obtained from the NEGF carrier densities $\rho_{n(p)}$ and current densities $J_{n(p)}$. Mobilities are then obtained from the relation $v_{n(p)} = \mu_{n(p)} F$. In order to minimize the influence of the contact resistances[8] (contained in the boundary self-energies) in the resulting carrier mobilities, the

[8]Notice that the carrier mobilities obtained from the average velocity in Eq. (4-12) correspond to "apparent mobilities", meaning that they contain the effect of both ballistic and diffusive processes (contact reflections are included in the ballistic component). In macroscopic devices, it is the diffusive component

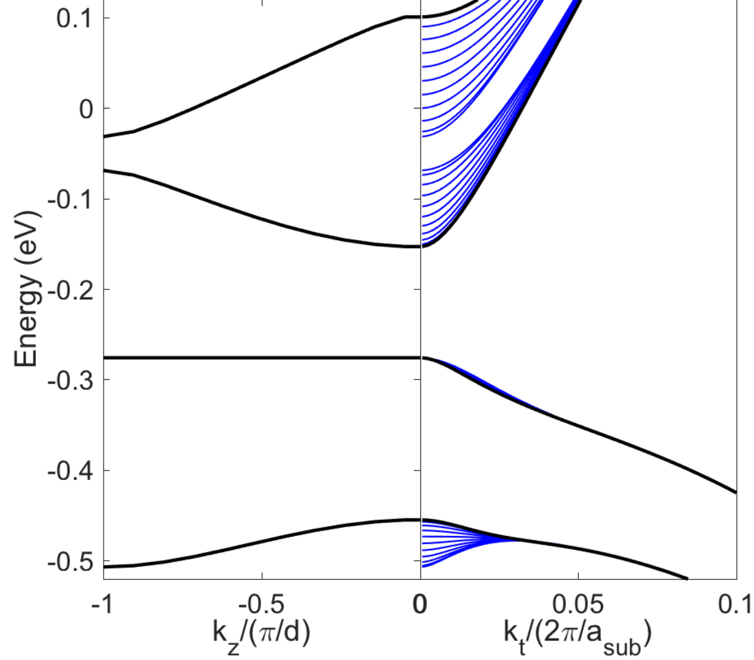


Figure 4-9: *The $\boldsymbol{k}\cdot\boldsymbol{p}$ subband structure of the SL considered in Ref. [100], plotted versus wavevectors along the in-plane (left panel) and growth directions (right panel). The in-plane dispersion for $k_z = 0$ is shown in black, while the blue curves are for equally spaced values of k_z up to the mini-Brillouin-zone boundary π/d (d is the SL period).*

average in Eq. (4-12) is computed for a section of 5 inner periods in the middle of a long SL sample (approximately 250 nm long).

The mobility study will be performed on a type-II SL consisting in the alternating sequence of 10.26 nm $InAs_{0.97}Sb_{0.03}$ and 1.89 nm $InAs_{0.55}Sb_{0.45}$, which was taken from Ref. [100]. Material parameters and Luttinger parameters are those from Tabs. **4-1** and **A-1**, respectively. The parameters of the alloys were obtained using linear interpolation of the parameters of the constituent materials, including also bowing effects according to Ref. [105]. Fig. **4-9** shows the electronic structure of the SL computed with Bloch boundary conditions, *i.e.*, the

of the mobility, also called "intrinsic mobility", the one of interest. As it was discussed in Ref. [104], the apparent mobility computed from the average velocity method gives a good approximation to the intrinsic mobility for sufficiently long samples. The length of the SL sample is more critical for electrons than for holes, since hole transport is usually dominated by diffusive processes, while electron transport is almost coherent for wide minibands. Our computations of the electron mobility were validated by comparing the results obtained from various SLs of increasing lengths. An alternative approach, which inherently neglects both the ballistic and contact effects, is the resistance scaling method, in which the intrinsic mobility is extracted from the dependence of the resistance on the device length, see [104].

eigenvalue problem $\mathbf{H}(\mathbf{k})\psi_j = E_j\mathbf{S}\psi_j$ is solved inserting [35]

$$\mathbf{F}_{\mathbf{k}}(z+d) = e^{\mathrm{i}k_z d}\mathbf{F}_{\mathbf{k}}(z), \tag{4-13}$$

in the boundary terms from Eqs. (2-107)-(2-108). Here, d is the length of one period of the SL, which for the present case is $d = 12.15$ nm, and k_z is the longitudinal wavevector introduced due to the structure periodicity in the growth direction. The cutoff wavelength[9] is found to be 11.8 μm, close to the experimental value of 12.5 μm extracted from low-temperature photoluminescence spectra [100]. In order to perform the NEGF simulation, we take a 12 period sample of the SL, discretized using the same parameters as in the previous structure, but with a reduced energy range ($E_{\text{min}} = -0.6$ eV, $E_{\text{max}} = 0.2$ eV). The contact Fermi levels are chosen such that an average hole density of $p_{\text{eq}} = 1 \times 10^{16}$ cm^{-3} is obtained at equilibrium (p-type SL), and a small bias voltage $V_{\text{bias}} = 1.1$ meV is then applied so the resulting electric field inside the structure is $F = 0.1$ kV/cm. A mode-space transformation is used to carry out the computations, using a total of $N_n = 250$ maximally localized basis functions.

The local density of states shown in Fig. **4-10**(a) reveals that the conduction miniband widths Δ_{c1} and Δ_{c2} are larger than the gap between them, while for the highest valence miniband, Δ_{v1} is smaller than 1 meV. This values are displayed in Fig. **4-10**(b), where we show a cut of the LDOS in Fig. **4-10**(a) at position $z = 76.3$ nm, although they can be obtained directly from the dispersion of the band structure along k_z, without the effect of homogeneous or inhomogeneous broadening. The difference in Δ, which in part is due to the relatively low conduction band offsets in InAs/InAsSb SLs, has a profound effect on the electron and hole transport properties. Although the spatially resolved spectral current in both n- and p-type SLs looks similar for small electric fields (see the continuous stripes crossing the band diagrams in Fig. **4-10**(c)), the nature of carrier transport for electrons and holes is very different. A closer inspection of the current components reveals that electron transport is mainly a coherent process through extended Bloch states at low fields, which can be expressed by the familiar Landauer-Büttiker formula

$$J_{\text{coherent}} = \frac{e}{\hbar}\int \frac{dE}{2\pi} T(E)\big[f_L(E) - f_R(E)\big], \tag{4-14}$$

where

$$T(E) = \frac{1}{A}\sum_{\mathbf{k}} \mathrm{Tr}\left[\mathbf{\Gamma}_L^B \mathbf{G}^R \mathbf{\Gamma}_R^B \mathbf{G}^A\right], \tag{4-15}$$

[9] In the context of photodetector devices, the cutoff wavelength refers to the maximum wavelength below which the device begins to exhibit significant absorption of incoming light. This value is determined by the bandgap of the material used in the device's absorption stage, and it can be estimated through the formula $\lambda_{\text{th}} = hc/E_g$.

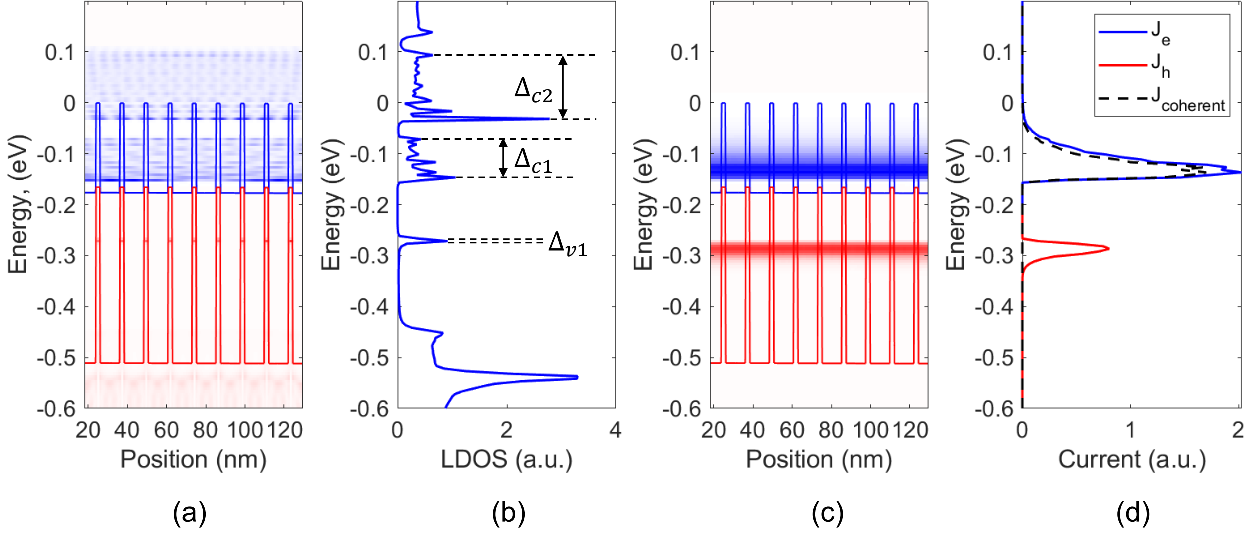


Figure 4-10: *(a) Local density of states for* $\boldsymbol{k} = 0$*, computed for a 12 period* $InAs_{0.97}Sb_{0.03}/InAs_{0.55}Sb_{0.45}$ *SL at* $T = 200$ *K. The applied electric field is* $F = 0.1$ *kV/cm. The first valence miniband (the sharp red peaks in the LDOS) is strongly localized in both energy and space, while the conduction minibands are much wider in energy. (b) Cut of the LDOS made at* $z = 76.3$ *nm, showing the spectral widths* $(\Delta_{c1,c2,v1})$ *of each miniband. (c) Spatially resolved current spectrum flowing through the SL. While electron transport is mostly a coherent process, hole transport is mainly assisted by phonon scattering. This is confirmed in (d) where a comparison is made between the spectral current at* $z = 76.3$ *nm (solid blue and red lines) and the coherent current coming from the Landauer-Büttiker formula (dashed black line).*

is the transmission probability, $\boldsymbol{\Gamma}^{B}_{L,R} = \mathrm{i}\left[\boldsymbol{\Sigma}^{RB}_{L,R} - \boldsymbol{\Sigma}^{AB}_{L,R}\right]$ are the broadening functions of the left and right contacts, and $f_{L,R}(E) = f(E - \mu_{L,R})$ are the corresponding occupation functions. On the other hand, hole transport is entirely noncoherent even in the low-field limit, meaning that the holes, strongly localized in the weakly coupled InAsSb quantum wells, can move across the miniband only if assisted by scattering mechanisms such as carrier-phonon interactions. This is in fact observed in Fig. **4-10**(d) where we compare the spectrally resolved current density at $z = 76.3$ nm (solid blue and red lines) with the coherent current from Eq. (4-14) (without the energy integration), J_{coherent} contributes most to the total electron current but its contribution to the hole current is negligible as most of it corresponds to scattering current.

In Fig. **4-11** we show the electron and hole drift velocities as a function of the electric field at $T = 200$ K. The velocities are almost linear for low-fields, confirming that the structure is long enough to reach the ohmic drift velocity regime. The critical field at which the electron drift velocity peaks is approximately 4 kV/cm, which corresponds to a potential drop per

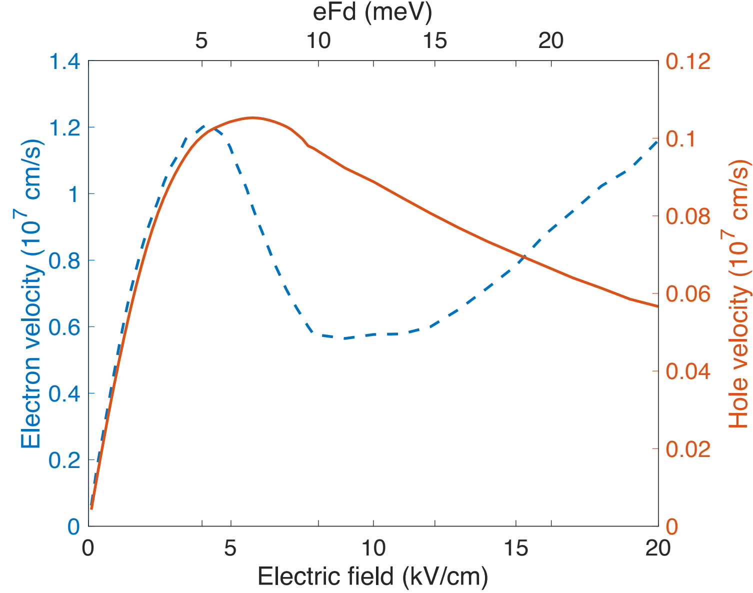


Figure 4-11: *Electron (left axis, dashed blue curve) and hole (right axis, solid red curve) velocities as a function of the applied electric field (lower axis) and of the potential drop per period eFd (upper axis), at $T = 200$ K.*

period $eFd = 5$ meV. From a semiclassical perspective, beyond this critical field, the electrons experience unscattered Bloch oscillations in the mini-Brillouin zone, which do not contribute to the net current. For higher fields, the Bloch oscillations should eventually lead to the breaking of the miniband into a ladder of Wannier-Stark (WS) levels, and carrier transport is best described by hopping between the states of the WS ladder rather than by miniband transport. However, due to the large width of the electron minibands and their relative small separation in energy, this condition is never reached. For higher fields, the electron velocity increases again because coherent tunneling between minibands (interminiband transport) becomes possible.

In n-type LWIR type-II SLs, a low hole velocity and therefore a low collection efficiency is expected from the large band-edge hole effective mass in the growth direction, and perhaps for this reason, most LWIR detectors have been based on p-type absorbers; the experimental evidence that hole mobility in type-II SLs is not as poor as expected was explained with the anisotropy and nonparabolicity of the valence band away from the band edge [106]. Our calculations indicate that the hole velocity, although approximately one order of magnitude smaller than the electron velocity, is significantly higher than expected. Esaki-Tsu theory [107] of superlattice transport in terms of Bloch oscillations fails for such a narrow miniband, as $eFd \gg \Delta_{v1}$ for any realistic field F. Hole transport is always dissipative. The spectral scattering rates indicate that near the flatband condition, the holes move across the superlattice by hopping up and down the miniband, emitting and absorbing polar optical

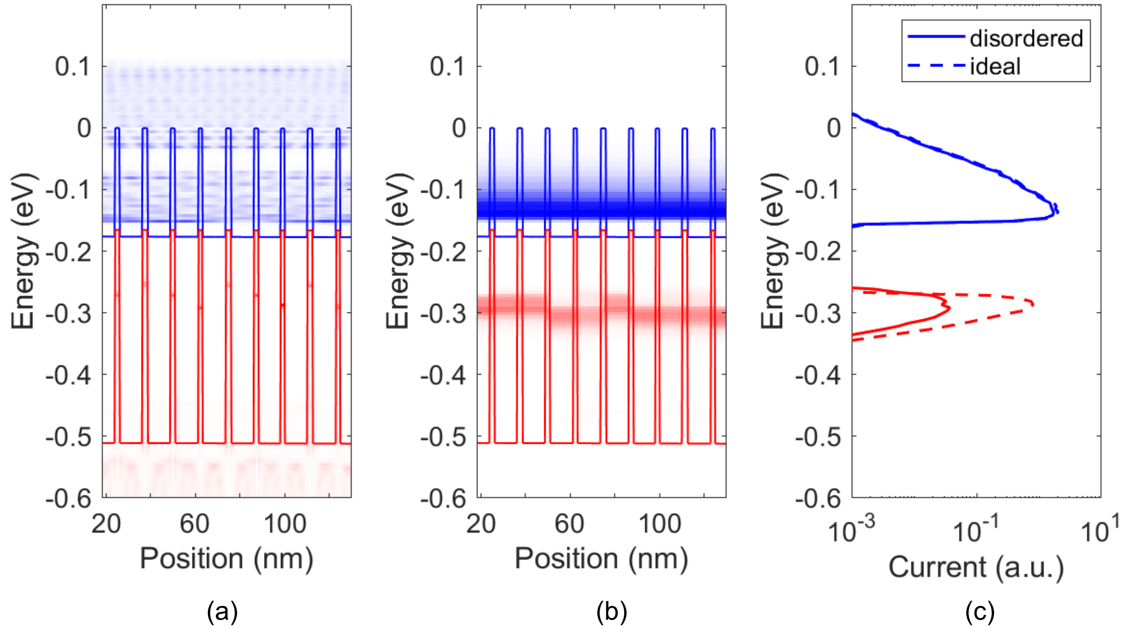


Figure 4-12: *(a) Local density of states for* $\boldsymbol{k} = 0$*, computed for the disordered SL. The applied electric field is* $F = 0.1$ *kV/cm. Energy alignments of the first valence miniband (sharp red peaks in the LDOS) are severely disrupted by the disorder, while the wider conduction minibands are not visibly affected, suggesting that hole transport is likely to be more sensitive to disorder. (b) This is confirmed by the spatially resolved current spectrum shown in the disordered structure. (c) Comparison of the spectral current at* $z = 76.3$ *nm for the ideal (dashed lines) and disordered (solid lines) case. The energy-integrated current drops by approx. one order of magnitude in the presence of geometrical fluctuations at* $T = 200$ *K.*

phonons. At higher fields, a WS ladder is clearly visible and electron hopping in the valence band is mostly a downhill process governed by the spontaneous emission of phonons[10].

The velocity-field curves in Fig. **4-11** are computed for an ideal SL, but the very concept of miniband becomes questionable when Δ becomes very small, due to crystal imperfections and other departures from ideality. Experimental investigations of vertical hole transport in type-II Sb-based SLs provide a clear indication of the presence of hopping mechanisms between trap states, with hole transport characteristics similar to those seen in amorphous semiconductors displaying Anderson localization. Olson *et al.* [100] estimated the hole mobility by measuring minority carrier lifetimes and the fraction of injected holes that diffuse

[10]Monte-Carlo calculations of hopping transport predict a weak temperature dependence of the drift velocity in these conditions, since phonon absorption and stimulated emission, whose rates increase with temperature, tend to cancel each other, and the only contribution to the current is therefore the spontaneous emissions of phonons. See Ref. [107]

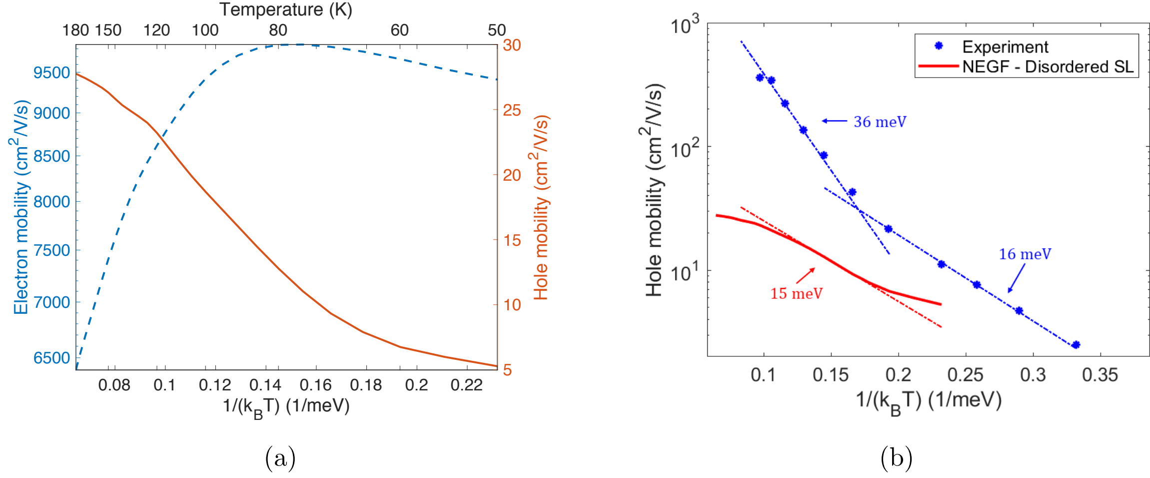


Figure 4-13: *(a) Electron (left axis, dashed blue curve) and hole (right axis, solid red curve) mobilities of the disordered SL as a function of inverse temperature, computed for an electric field of 0.1 kV/cm. (b) Comparison of the computed hole mobility with the experimental results reported by Olson et al. in Ref. [100].*

in the T2SL base region of a bipolar transistor[11]; different slopes in the Arrhenius plot of the hole mobility were attributed by the authors to distinct transport regimes, from miniband transport of thermally activated carriers, to a regime dominated by hopping within defect states above the miniband mobility edge. In a SL, the disorder may originate from crystal defects, interface roughness, and compositional and thickness fluctuations of the layers. In order to investigate carrier transport in disordered SLs, we randomly add or subtract one ML to each layer, breaking the translational symmetry along the growth axis. The effect of the disorder is particularly evident for the hole miniband, which appears to be severely disrupted; see the local density of states in Fig. **4-12**(a) and the spectral current densities in Fig. **4-12**(b). And, in fact, the hole current at $T = 200$ K is reduced by one order of magnitude with respect to the ordered structure, while the electron current is nearly unaffected at the same temperature, see Fig. **4-12**(c).

In Fig. 4.13(a) we show the electron and hole low-field mobilities (at $F = 0.1$ kV/cm) of the disordered SL as a function of inverse temperature. The electron mobility μ_n shows the conventional phonon-limited temperature dependence above 80 K, in which carrier-phonon scattering slows down the electron velocity, which is typical of miniband conduction. At lower

[11]The vertical mobility of minority carriers can also be extracted from the electron beam induced current (EBIC) technique and time-resolved photoluminescence (TRPL) measurements, combined with analytical models for the generation of the carriers and their subsequent collection at the contacts. Experimental measurements of vertical superlattice mobilities are usually much more challenging than transverse mobilities experiments owing to the complexity of the active region, the finite size of the electron beam, and the presence of surface recombination processes, see [108].

temperatures, the energy-resolved current density becomes spectrally narrower, involving mainly the less extended states near the bottom of the conduction miniband, which results in a slight decrease of μ_n. On the other hand, the hole mobility is thermally activated in the whole temperature range. The calculated hole mobilities are smaller than the experimental results obtained by Olson (see Fig. 4.13(b)), especially at higher temperatures, probably because the disorder is applied to the whole SL, irrespectively of the lateral degree of freedom. In general, in a SL sample of macroscopic size, the well and barrier thicknesses may fluctuate in the layer plane, meaning that fully localized states along the growth axis may coexist with miniband states sufficiently extended to allow transport, so that percolative transport along high conduction paths may even conceal the presence of localized states from mobility experiments.

The activation energies associated with the different slopes observed in Fig. 4.13(b) can be computed by assuming a relation between the temperature and the mobility of the form

$$\mu_h \propto e^{\frac{-E_a}{k_B T}}. \tag{4-16}$$

The estimated value from our calculations $E_a = 15$ meV is very close to that reported in Ref. [100] for region 3 ($E_a = 16$ meV), in which transport was considered to be dominated by hopping between localized states near the mobility edge, in the exponential tail of the valence miniband. As it was discussed in Ref. [109], this values are related to the average energy needed for a trapped hole in order to reach a neighbouring SL layer.

4.5 Transport across a barrier in a superlattice

Among possible candidates for the assessment of the sophisticated physics in type-II SLs, barrier infrared detectors provide the opportunity to analyze separately carrier transport properties of electrons (minority carriers in pBp configurations) and holes (minority carriers in nBn configurations) [110, 111]. Moreover, this structures represent an important part in IR absorbers like the ICIP studied in Sec. 4.3, acting also as relaxation phases and facilitating inter-miniband tunneling. In this section we consider an nBn detector, consisting of an undoped 8 period GaSb/AlSb (18/24 Å) superlattice barrier interposed between a 40 period InAs/GaSb (24/18 Å) absorber, with n-type 10^{17} cm^{-3} doping. The same superlattice is used also for the top and bottom contact regions, where the doping is increased to 5×10^{18} cm^{-3}. An schematic representation of the structure is shown in Fig. **4-14**.

4.5.1 Numerical considerations

The simulation of this structure is carried out in a similar fashion as before, using uniform grids in energy, space and transverse momentum, with $\Delta E = 5$ meV, $\Delta z = 0.5$ nm,

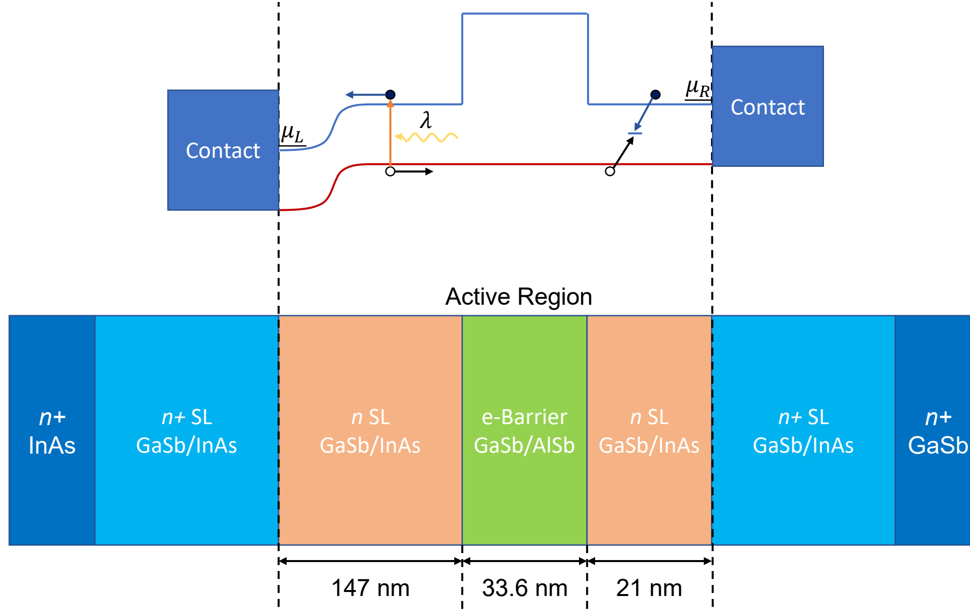


Figure 4-14: Schematic of the nBn superlattice absorber studied in this section. The active region is composed of an 8-period GaSb/AlSb electron barrier sandwiched between a 40-period InAs/GaSb superlattice absorber. An illustration of the operation of the device is shown above. Photogenerated minority holes diffuse thorough the barrier and reach the right contact layer where they recombine (*e.g.*, through an intermediate defect state by SRH).

$k_{\max} = 2.0$ nm^{-1}, $E_{\min} = -1.0$ eV and $E_{\max} = 0.8$ eV. For simplicity, we include only the first conduction and heavy-hole bands in the effective mass approximation, with the non-interacting Hamiltonian given by Eq. (2-102). Furthermore, instead of using a mode-space representation, we follow a different route here by computing the scattering self-energies according to the Büttiker formalism, using Eqs. (2-210) and (2-212) for the acoustic and optical deformation potential self-energies, respectively. The main advantage of the Büttiker probes self-energies is that they only depend on the diagonal of the (covariant) GFs matrices, and therefore only these diagonals need to be stored on each iteration of the inner-loop, reducing significantly the memory usage. Additionally, the use of the Büttiker formalism combined with an effective mass approximation of the dispersion relation result in a very convenient property for the retarded component of the GF, as it only depends on the energy E and the norm of the transversal wavevector $k = |\mathbf{k}|$ through the relation [58]

$$\mathbf{G}^R(k, E) = \mathbf{G}^R(0, E - E_k), \tag{4-17}$$

with $E_k = \hbar^2 k^2 / 2m_t^*$ and m_t^* the transversal effective mass. This property allows to perform analytically integrations over the transversal momentum by making the substitution $E' =$

$E - E_k$. For example, the acoustic phonon self-energy (2-210) can be computed as[12]

$$\begin{aligned}\Sigma^R_{\text{aco},ii}(E) = \frac{D_a^2 k_B T}{A\rho u_l^2 a_L}\sum_{\mathbf{k}'}\tilde{G}^R_{i,i}(0,E-E_{k'}) &\approx \frac{D_a^2 k_B T}{2\pi\rho u_l^2 a_L}\int_0^{k_{\max}} dk'\,k'\,\tilde{G}^R_{i,i}(0,E-E_{k'}) \\ &= \frac{D_a^2 k_B T m_t^*}{2\pi\rho u_l^2 a_L \hbar^2}\int_0^{E_{k_{\max}}} dE_{k'}\,\tilde{G}^R_{i,i}(0,E-E_{k'}) \\ &= \frac{D_a^2 k_B T m_t^*}{2\pi\rho u_l^2 a_L \hbar^2}\int_{E-E_{k_{\max}}}^{E} dE'\,\tilde{G}^R_{i,i}(0,E'),\end{aligned} \tag{4-18}$$

where $\tilde{\mathbf{G}}^R = \mathbf{S}\mathbf{G}^R\mathbf{S}$ is the covariant representation of the retarded GF and $E_{k_{\max}} = \hbar^2 k_{\max}^2/2m_t^*$. Similarly, for the optical phonon self-energy (2-212) we obtain

$$\begin{aligned}\Sigma^R_{\text{op},ii}(E) = \frac{\hbar D_{tK}^2}{2A\rho\omega_0 a_L}&\sum_{\mathbf{k}'}\Big\{\Big[n_0 - f_{\text{FD}}\Big(E - \hbar\omega_0 - E_{\text{BP}}^{(m)}\Big) + 1\Big]\,\tilde{G}^R_{i,i}(0,E-E_{k'}-\hbar\omega_0) \\ &+ \Big[n_0 + f_{\text{FD}}\Big(E + \hbar\omega_0 - E_{\text{BP}}^{(m)}\Big)\Big]\,\tilde{G}^R_{i,i}(0,E-E_{k'}+\hbar\omega_0)\Big\} \\ &\approx \frac{D_{tK}^2 m_t^*}{4\pi\rho\hbar\omega_0 a_L}\Big\{\Big[n_0 - f_{\text{FD}}\Big(E - \hbar\omega_0 - E_{\text{BP}}^{(m)}\Big) + 1\Big]\int_{E-E_{k_{\max}}}^{E} dE'\,\tilde{G}^R_{i,i}(0,E'-\hbar\omega_0) \\ &+ \Big[n_0 + f_{\text{FD}}\Big(E + \hbar\omega_0 - E_{\text{BP}}^{(m)}\Big)\Big]\int_{E-E_{k_{\max}}}^{E} dE'\,\tilde{G}^R_{i,i}(0,E'+\hbar\omega_0)\Big\}.\end{aligned} \tag{4-19}$$

Since the Büttiker probes self-energies depend $\mathbf{G}^R$ alone, we can avoid solving the Keldysh equation (2-76) inside the inner-loop iterations, and due to the property (4-17), just $\mathbf{G}^R(k=0,E)$ is required. The sum over transverse wavevector of the lesser GF, necessary for the calculation of the observables, is carried out as follows

$$\begin{aligned}\frac{1}{A}\sum_{\mathbf{k}}\mathbf{G}^<(k,E) &= \frac{1}{A}\sum_{\mathbf{k}}\mathbf{G}^R(k,E)\boldsymbol{\Sigma}^<(k,E)\mathbf{G}^A(k,E) \\ &\approx \frac{m_t^*}{2\pi\hbar^2}\int_0^{E_{k_{\max}}} dE_k\,\mathbf{G}^R(0,E-E_k)\boldsymbol{\Sigma}^<(k,E)\mathbf{G}^A(0,E-E_k).\end{aligned} \tag{4-20}$$

with $\boldsymbol{\Sigma}^<(k,E) = \mathrm{i}f_{\text{FD}}\big(E - E_{\text{BP}}^{(m)}\big)\boldsymbol{\Gamma}(k,E)$, and $\boldsymbol{\Gamma}(k,E) = \mathrm{i}\big[\boldsymbol{\Sigma}^R(k,E) - \boldsymbol{\Sigma}^A(k,E)\big]$. Here $\boldsymbol{\Sigma}^<(k,E) = \boldsymbol{\Sigma}^{<B}(k,E) + \boldsymbol{\Sigma}^<_{\text{aco}}(E) + \boldsymbol{\Sigma}^<_{\text{op}}(E)$, and notice that the retarded boundary self-energy from Eq. (2-132) satisfies the property (4-17), *i.e.*, $\boldsymbol{\Sigma}^{RB}(k,E) = \boldsymbol{\Sigma}^{RB}(0,E-E_k)$. Scattering processes related to optical and SRH interband transitions will be also taken into account in the simulation. Since the electron-phonon and SRH self-energies are not compatible with the property (4-17) allowing the analytic integration over transverse wavevectors,

[12] Recall that Σ^R_{aco} and Σ^R_{op} are independent of the transverse momentum.

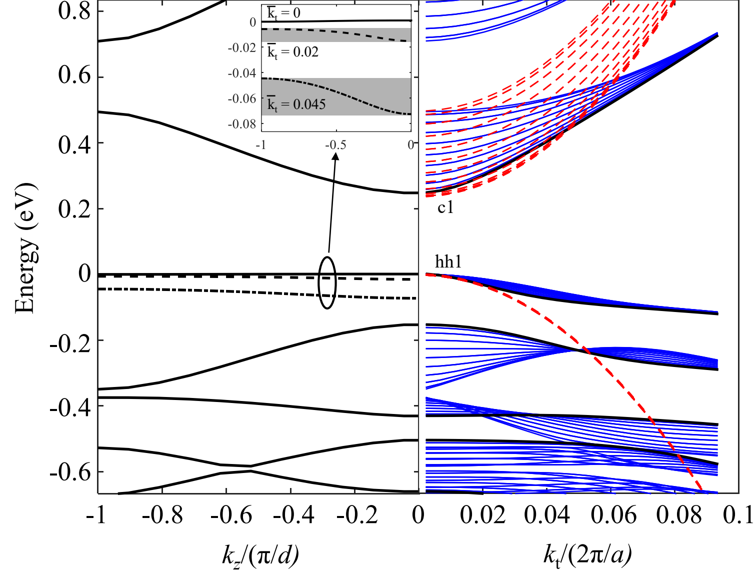


Figure 4-15: Subband structure of a superlattice consisting in the alternating sequence of a 24/18 Å InAs/GaSb layers, computed with a 8-band multiband $\mathbf{k} \cdot \mathbf{p}$ model for wavevectors along the growth (left panel) and in-plane (right panel) directions. The in-plane dispersion for $k_z = 0$ is shown in black, while the blue curves are for equally spaced values of k_z up to the mini-Brillouin-zone boundary π/d (d is the superlattice period). The inset in the left panel shows the first heavy-hole (hh1) subband dispersion along k_z for different values of the (normalized) transverse wavevector k_t; the coloured strips mark the corresponding minibands.

we take a different route by computing the net interband recombination rates with the semi-classical expressions. In order to investigate the extraction efficiency of type-II SL detectors, we assume an electron recombination rate of the form

$$R_n(z) = R_{\text{SRH}}(z) - G_{\text{opt}}(z), \tag{4-21}$$

where G_{opt} is a constant optical generation rate computed from the absorption in the material and the photon flux, and R_{SRH} is the (net) SRH recombination rate given by Eq. (3-53) (n and p are the NEGF electron and hole densities in Eqs. (2-82) and (2-83), respectively). The net recombination rate is then inserted in the right hand side of the NEGF continuity equation, Eq. (2-56), which reads in our finite-element representation

$$\frac{\partial J_n}{\partial z}(z_i) = \frac{1}{A\Delta z} \sum_{\mathbf{k}} \int \frac{dE}{2\pi\hbar} \left\{ \bar{\mathbf{H}} \mathbf{G}^<(0, E - E_k) - \mathbf{G}^<(0, E - E_k) \bar{\mathbf{H}} \right\}_{i,i} = R_n(z_i), \tag{4-22}$$

with $\bar{\mathbf{H}} = E\mathbf{S} - \mathbf{H}$, and the integration is over the conduction miniband. Similarly the hole continuity equation is solved by taking $R_p = -R_n$ and integrating over the valence miniband.

The solution of Eq. (4-22) is obtained by adjusting the Büttiker probes Fermi-levels using a fast converging Newton-Raphson algorithm (both the Büttiker self-energies and R_n are set to zero at the contacts). We also include space-charge effects through the solution of the Poisson equation. The simulations were performed at dark and under the illumination of a monochromatic source with intensity $I_\gamma = 0.1$ W/m^2. The influence of the contact regions ($n+$ SL at both sides of the structure in Fig. **4-14**) is folded into the boundary self-energy according to Eqs. (2-117a),(2-117b),(2-118a) and (2-118b).

Table 4-2: List of simulation parameters.

Quantity	Value	Unit
D_a	8	eV
D_{tK}	1×10^9	eV/cm
ρ	5.32	g/cm^3
u_l	6.56	km/s
$\hbar\omega_0$	30	meV
τ_n	10^{-8}	s
τ_p	10^{-8}	s
$m^*_{e,t}$	0.045	m_0
$m^*_{e,z}$	0.035	m_0
$m^*_{h,t}$	0.05	m_0
$m^*_{h,z}$	0.28	m_0

Finally, the band parameters (effective masses) of the two-band effective mass model are computed from the subband structure of the SL absorber. Fig. **4-15** shows the subband structure of the superlattice absorber computed with an 8-band $\mathbf{k} \cdot \mathbf{p}$ model and Bloch boundary conditions from Eq. (4-13). The calculated cutoff wavelength of the absorber is 4.8 μm. While, the lowest conduction sub-band (c1) shows strong dispersion along both growth (z) and in-plane (x, y) directions, the highest valence sub-band (hh1) is highly anisotropic, being actually 'light' in the in-plane directions, and almost dispersion-less along the growth direction. We now proceed to approximate the electronic structure with parabolic valleys. Specifically, we treat the longitudinal effective masses $m^*_{e,z}$, $m^*_{h,z}$ as fitting parameters to reproduce the position of the c1 and hh1 mini-bands at zero transverse momentum, which gives the effective bandgap of the SL, and then we choose the transverse effective masses $m^*_{e,t}$, $m^*_{h,t}$ so as to fit the in-plane dispersion. The effective mass approximation (red dashed lines in Fig. **4-15**) reproduces well the $\mathbf{k} \cdot \mathbf{p}$ electronic structure at small wave vectors. The nonparabolicity of the c1 sub-band due to the interaction between conduction and valence bands, and the slight increase in the dispersion of the hh1 sub-band along the growth direction due to heavy-hole and light-hole mixing at intermediate values of the in-

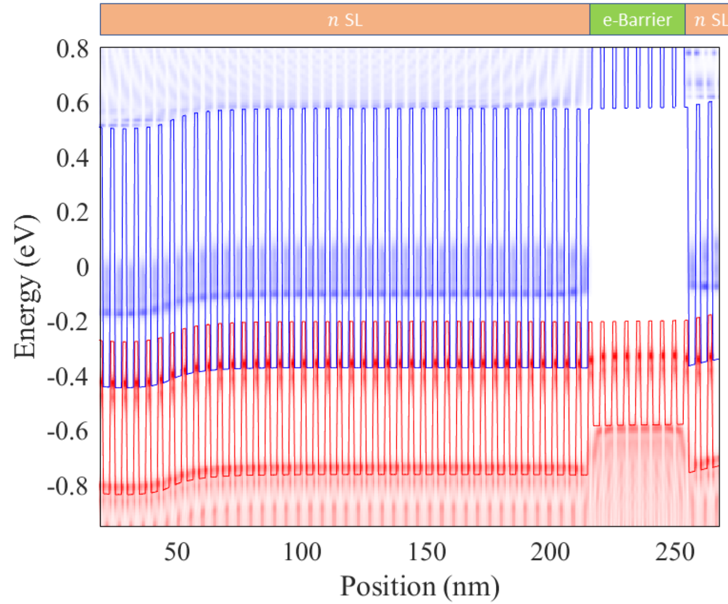


Figure 4-16: Local density of states of the nBn structure, shown for $\mathbf{k} = 0$ (color maps). The reverse bias applied to the top (right) contact is 0.1 V. The barrier is implemented with an undoped GaSb/AlSb superlattice to avoid the valence band offset that typically arises when AlGaSb is used as the barrier material [113]. The first conduction miniband of the superlattice absorber has an spectral width of approximately 0.2 eV, while the first valence miniband is strongly localized in the weakly coupled GaSb quantum wells. All simulations are performed at 150 K.

plane momentum k_t (see inset), cannot be captured by single-band models. The last feature has implications in hole transport, as discussed in [112]. The effective masses fitting the sub-band structure of the absorber layer are reported in Tab. **4-2**.

4.5.2 Simulation results

Fig. **4-16** shows the LDOS (color maps), and the conduction and valence band edges E_C, E_V (blue and red thin lines) of the nBn structure, computed with the NEGF approach at a reverse bias of 0.1 V (applied to the right contact) and at $T = 150$ K. As expected, the valence miniband is narrower than the conduction miniband. The states in the GaSb/AlSb barrier are not visible in Fig. **4-16**, as the energy scale has been limited to 0.8 eV (above this energy the spectral current is negligible). The different widths of the electron and hole minibands has important implications on carrier transport: electron transport is mainly a coherent process through extended Bloch states, while hole transport is entirely non-coherent, meaning that the holes are strongly localized in the weakly coupled quantum wells, and their motion across the miniband is made possible by carrier-phonon scattering. The possible presence of

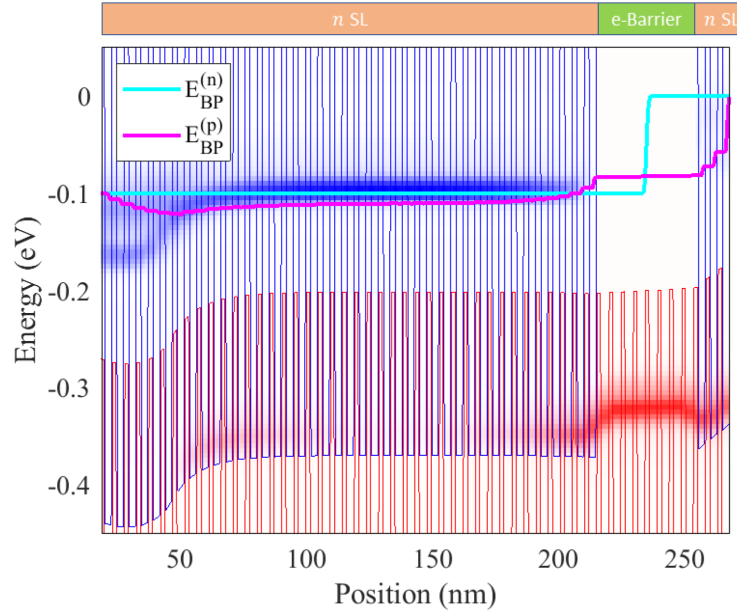


Figure 4-17: Spatially- and spectrally-resolved current density (color maps) with a light illumination intensity of $I_\gamma = 0.1\,\mathrm{W/m^2}$, and at a reverse bias of 0.1 V. The Fermi levels of the Büttiker probes (NEGF) and the quasi-Fermi levels of the carriers (SPDD) are shown in solid and dashed lines, respectively.

compositional and geometrical fluctuations, not considered here, could make the difference between electron and hole mobilities even more striking [91].

The spectral current computed under illumination at $T = 150$ K is shown in Fig. **4-17**. The narrow strips following the miniband edges represent electrons and holes photogenerated in the absorber diffusing in opposite directions, the electrons to the left, the holes towards the contact layer beyond the electron barrier. The spreading of the electron spectral current in the highly doped end of the absorber is indicative of electron relaxation by phonon scattering. The Fermi levels of the Büttiker probes can be also observed in Fig. **4-17**. In the barrier region, the electron Fermi level has an error-function shape, almost step-like in the NEGF calculations, *i.e.*, the electrons tunneling from the top contact through the barrier, rapidly become members of the absorber 'ensemble' adjusting to the Fermi level on the other side of the barrier (a similar behavior can be observed in short-channel FETs [114]). In the almost field-free superlattice absorber, the NEGF Fermi level of the (majority) electrons is almost constant. On the other hand, the Fermi level of the (minority) holes show a stair-case profile, which is more visible near the barrier layer, where the electric field is present. This behavior suggests that hole transport is best described by sequential tunneling, *i.e.*, tunneling to the neighboring well followed by relaxation.

The energy-integrated electron and hole currents under illumination computed with NEGF

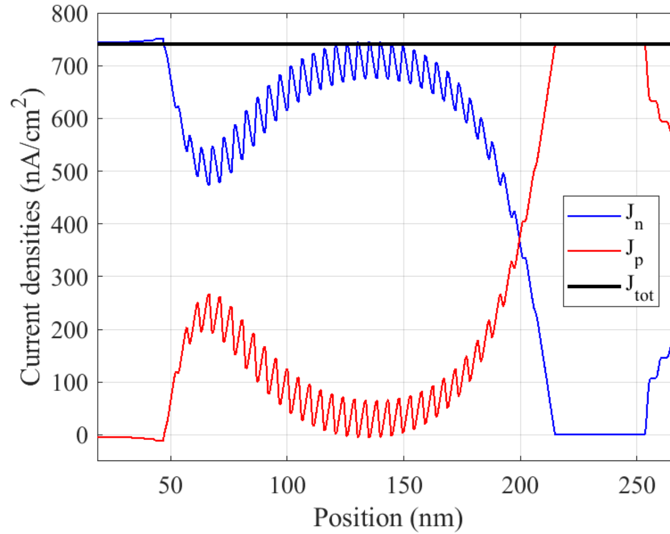


Figure 4-18: Energy-integrated spatial profile of electron and hole currents obtained from NEGF (blue and red solid lines). As GR processes are suppressed in the barrier (see Fig. **4-19**), the electron and hole currents remain constant in this region. The total NEGF current (solid black line) is perfectly conserved.

is shown in Fig. **4-18**. In our calculations, current conservation is achieved in just few iterations of the Newton algorithm. This is perhaps the most useful advantage of the Büttiker formalism. In the superlattice absorber, the electron and hole currents have the typical shape expected in a selectively contacted device. In contrast to p-n junction photodiodes, in nBn detectors, the photogenerated holes crossing the barrier enter the contact layer as minority carriers, where they recombine at a rate governed by the carrier lifetime, prolonging the duration of the photoevent [110]. The majority carriers are effectively blocked by the barrier, as the energy-integrated electron current is virtually zero in the barrier region. The carrier extraction efficiency

$$\eta_{\mathrm{ext}} = \frac{J_{\mathrm{SC}}}{J_{\mathrm{gen}}}, \tag{4-23}$$

defined as the ratio between the short-circuit current J_{SC} and the generation current J_{gen}, given by Eq. (4-11), is found to be limited to 25% by the short diffusion length of the holes. Finally, the spatial profile of the SRH recombination rate computed in dark and in illumination conditions confirms that SRH generation is mostly limited to the depletion region on the left side of the electron barrier, see Fig. **4-19**. Trap-assisted tunneling (TAT) is already included in both NEGF formulations. Field-enhancement factors, which are introduced in the expression of the SRH rate to account for the presence of the tail states that participate to the recombination process, as was already discussed in Sec. 3.3, are not needed here, since this effect is already included in the quantum NEGF densities.

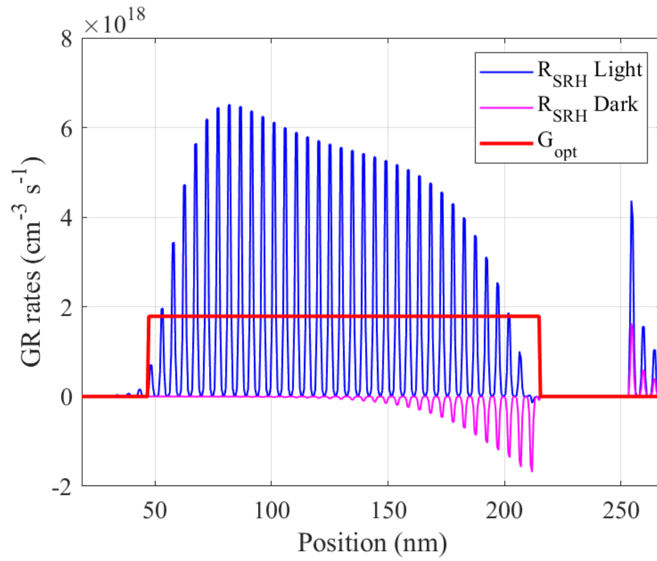


Figure 4-19: Spatial profile of the SRH recombination rate in dark and illumination conditions. For simplicity, the optical generation is assumed constant in the absorbing region. The competition of SRH recombination and optical generation in the superlattice absorber is clearly visible.

Chapter 5

Conclusions and Outlook

In this thesis, we have presented a theoretical investigation of carrier transport in optoelectronic devices by means of a rigorous quantum-kinetic model. The theoretical framework offered by the nonequilibrium Green's function formalism allows to incorporate coherent transport, quantization, and microscopic particle interactions on equal footing, resulting in a powerful method for the simulation of nanostructured devices dominated by quantum effects. Computer-aided design (CAD) tools for optoelectronic devices commonly produce standard outputs such as carrier densities, current densities, scattering rates, and emission spectra. In this work, these quantities are obtained through the nonequilibrium statistical averaging of the device's Green's function and self-energies, resolved in both space and energy. Moreover, NEGF provides a reliable approach to the modeling of realistic devices, since it only relies on fundamental parameters characterizing the geometrical, electrical and optical properties of the constituent bulk materials. We complement our model with a multiband description of the electronic structure based on the $\mathbf{k}\cdot\mathbf{p}$ theory, which is essential for the correct representation of interband tunneling and optical transitions. However, the high degree of accuracy achieved by this approach comes at the cost of a high computational burden, which make it unpractical to simulate structures longer than few hundreds of nanometers.

Our in-house implementation of the NEGF equations, based on a finite-element discretization of the longitudinal space, has been applied in the context of GaN-based LEDs and superlattice infrared photodetectors. Both are relevant structures subject to a high degree of investigation at present, the former for blue and UV light emitters, and the latter for its versatility and convenience with respect to state-of-the-art infrared imaging technologies (*e.g.*, detectors based on bulk HgCdTe).

First, in chapter 3 we have presented a study of trap-assisted tunneling in GaN-based LEDs. The first step towards a quantum-kinetic description of defect recombination assisted by tunneling with NEGF is the determination of the associated scattering self-energy, for which we have provided a theoretical derivation starting from the perturbation expansion the device

Green's function, and assuming that the isolated defect states interact with the extended states of the device by multiphonon emission. The multiband NEGF formulation of defect-mediated scattering is a novel result of this work. The first attempt toward a quantum kinetic description of Shockley-Read-Hall recombination with NEGF was made by *U. Aeberhard* in [65] using separated representations for the conduction and valence band Green's functions. Starting from this work, we have developed a more general theory adapted to our multiband $\mathbf{k} \cdot \mathbf{p}$ implementation. Starting from the principle of detailed balance of the capture and emission rates at the defect level, we have demonstrated that the net capture rate of electrons in the conduction band is equivalent to the semiclassical formulation from Shockley-Read-Hall, when the Green's function is assumed to satisfy quasi-equilibrium conditions. Here, the defect self-energy has been computed with three models: optical deformation potential, polar optical (Fröhlich model) and multiphonon theory, resulting in a good approximation of the semiclassical result in all three cases, and most importantly, they all give the same slope which characterizes the zero-field SRH recombination process (see Fig. **3-2**).

Proceeding further, we have studied the enhancement of defect recombination due to tail state formation, when the structure is subjected to high electric fields. As a first approximation, we have considered a GaN *p-n* diode with highly doped layers, with a single midgap defect level placed at the junction position where the electric field is maximum, observing the resulting recombination current. By separating the contributions of the SRH self-energy into direct and tunneling assisted components, we have demonstrated that the recombination rate obtained from the direct component is able to reproduce the results expected in a semiclassical situation (drift-diffusion simulation with the characteristic ideality factor of 2), while the full self-energy gives enhanced recombination with an ideality factor of 5.5 (see Fig. **3-8**). This result establishes a connection between NEGF and semiclassical SRH recombination models, and it also shows that the main contribution of the recombination current, giving rise to such high values of the ideality factor, comes from defect capture assisted by the tunneling of carriers into subgap states. Then, we moved to a more realistic case study: single quantum-well InGaN/GaN LED in the subthreshold forward bias regime. Experimental data taken from the literature [12] was used to contrast the simulation results. The computed ideality factors are in agreement with experimental data (see Fig. **3-15**). Our calculations show that the standard SRH formula, computed with the correct quantum carrier densities, can reproduce NEGF results (see Fig. **3-14**). This result suggests that trap-assisted tunneling can be described with drift-diffusion solvers complemented with appropriate quantum corrections for the calculation of the local density of states. Carrier densities may be estimated, *e.g.*, from localization landscape approaches, or more rigorously from the full eigendecomposition of the Schrödinger equation. Among possible quantum-corrected semiclassical approaches to study carrier transport in LEDs, the Schrödinger-Poisson drift-diffusion model seems promising. DD models, complemented with a quantum-corrected LDOS, may be the only viable approach to the numerical simu-

lation of complex realistic structures. Additional work is needed to verify the accuracy of such quantum-corrected semiclassical models, especially in the high-injection regime, where out-of-equilibrium phenomena such as hot carrier transport become relevant.

In chapter 4 we have moved to the simulation of carrier transport in realistic infrared detectors based on type-II superlattice absorbers. We have started by describing a mode-space approach in which the NEGF equations are projected into a lower-order basis set of maximally localized problem-matched basis functions. This approach enables the multiband simulation of longer structures by reducing the memory requirements of the NEGF operators that need to be stored on each iteration of the self-consistent loop. Unlike other projection-based approaches, we propose an additional self-energy term taking into account the contribution of the remote basis set and allowing to conserve the total current. We have demonstrated the robustness of the method by modeling an interband cascade infrared photodetector (ICIP) of nearly 150 nm long using an 8-band $\mathbf{k} \cdot \mathbf{p}$ description of the electronic dispersion. An inspection of the spectrally resolved current density reveals the presence of miniband transport in the superlattice absorber, where current is generated due to the absorption of carriers (see Fig. **4-6**). The optically generated electron current is then transformed into a hole current aided by the quantum-engineered hole and electron barriers that complement the device, and which serve as relaxation and tunneling phases (see Fig. 4.7(b)). The crucial role of phonon relaxation can be also observed in the spectral scattering current plot, assisting the cascade of the minority electrons from the conduction miniband of the superlattice absorber across the graded superlattice acting as relaxation region. Overall, the mode-space approach, corrected to include the effect of the remote bands, is revealed to be an enabling tool towards the simulation of realistic T2SL photodetectors.

We proceeded further by performing a mobility study in an $InAs_{0.97}Sb_{0.03}/InAs_{0.55}Sb_{0.45}$ superlattice. Although macroscopic quantities such as carrier mobilities and lifetimes are admittedly not germane to the NEGF formalism, they represent critical ingredients needed in quantum-corrected drift-diffusion models which can be investigated with genuine quantum transport models to obtain a semiclassical perspective of the underlying phenomena. Our results show an almost linear low-field trend of the electron and hole velocities at $T = 200$ K (see Fig. **4-11**), up to a critical value of approximately 4 kV/cm. At high fields, it is expected that the minibands break into a Wannier-Stark ladder, where carrier transport occurs mainly by hopping assisted by phonon emission. In fact, the small spectral width of the first heavy-hole miniband HH1 ($\Delta_{v1} < 1$ meV) contributed to its subsequent breakdown even in the low-field limit, making hole transport highly dependent on the population of phonons assisting the hopping transitions. On the other hand, electron transport was observed to remain nearly unaffected by this effect, as most of the electron current corresponded to coherent propagation (see Fig. **4-10**). In reality, disorder effects due to compositional or layer fluctuation can greatly impact the transport in superlattice structures. This was

investigated by introducing random fluctuations in the layers width, which severely disrupted the hole miniband alignment (see Fig. **4-12**). An inspection of the electron and hole low-field mobility over a wide range of temperatures (see Fig. 4.13(a)) revealed that μ_n displays a conventional phonon-limited dependence above 80 K, while μ_p is thermally activated in the whole energy range, mainly because of the dependence of hole transport in the phonon population assisting the hopping process. The hopping activation energy estimated from our calculations was $E_a = 15$ meV.

The last case study considered in this work consisted in a barrier IR detector made with type-II superlattices, specifically an nBn detector. This structures are of particular importance from an analytical point of view, because they provide the opportunity to analyze separately carrier transport of electrons and holes. In this case we have moved from the analysis of a single section of a device to the description of a full device using NEGF, including also space-charge effects and SRH recombination, which is a novel result. Compared with the previous analyses, we have followed a different route by including phonon scattering within the Büttiker probes formalism, in which local self-energies depend only on the diagonal of the Green's functions, hence reducing the amount of quantities that must be stored on each iteration. Under illumination conditions, we have observed that the barrier was able to successfully block the majority carrier current, while the minority carrier current rapidly recombined after crossing the barrier (see Fig. **4-18**). The carrier extraction efficiency of the structure was found to be 25%.

In conclusion, we have presented a numerical investigation of carrier transport based on a solid theoretical background, which represents a reliable approach to the simulation of nanostructures with arbitrary band alignment and bias/temperature conditions. In this view, we think that the most promising strategy for the simulation of large devices featuring nanostructure is the use of multiscale approaches, by integrating NEGF simulations of small regions (obtaining, *e.g.*, the tunneling currents, quantum charge, effective mobilities, etc) into a semiclassical simulation for the complete device. In the future, we expect to move towards this direction, and also to investigate more complex scattering processes like impact ionization, or (trap-assisted) Auger processes, which require investigation at a fundamental level.

Appendix A

The 8-band $\mathbf{k}\cdot\mathbf{p}$ Luttinger-Kohn Hamiltonian

A.1 Zinc-blende model

The starting point of the finite-element discretization of the Dyson and Keldysh equations is the eight-band Pidgeon-Brown or Enders $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian [115] for strained tetrahedral semiconductors

$$\mathbf{H}_{\mathrm{Bulk}}^{k\cdot p}(\underline{\mathbf{k}}) = \mathbf{H}_0 + \mathbf{H}_\varepsilon + \mathbf{H}_1(\underline{\mathbf{k}}) + \mathbf{H}_2(\underline{\mathbf{k}}), \tag{A-1}$$

in the zone-center basis

$$\{S\uparrow, X\uparrow, Y\uparrow, Z\uparrow, S\downarrow, X\downarrow, Y\downarrow, Z\downarrow\},$$

where

$$\mathbf{H}_0 = \begin{bmatrix} E_g & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\Delta/3 & -\mathrm{i}\Delta/3 & 0 & 0 & 0 & 0 & \Delta/3 \\ 0 & \mathrm{i}\Delta/3 & -\Delta/3 & 0 & 0 & 0 & 0 & -\mathrm{i}\Delta/3 \\ 0 & 0 & 0 & -\Delta/3 & 0 & -\Delta/3 & \mathrm{i}\Delta/3 & 0 \\ 0 & 0 & 0 & 0 & E_g & 0 & 0 & 0 \\ 0 & 0 & 0 & -\Delta/3 & 0 & -\Delta/3 & \mathrm{i}\Delta/3 & 0 \\ 0 & 0 & 0 & -\mathrm{i}\Delta/3 & 0 & -\mathrm{i}\Delta/3 & -\Delta/3 & 0 \\ 0 & \Delta/3 & \mathrm{i}\Delta/3 & 0 & 0 & 0 & 0 & -\Delta/3 \end{bmatrix}, \tag{A-2}$$

$$\mathbf{H}_1(\underline{\mathbf{k}}) = \begin{bmatrix} 0 & \mathrm{i}Pk_x & \mathrm{i}Pk_y & \mathrm{i}Pk_z & 0 & 0 & 0 & 0 \\ -\mathrm{i}Pk_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\mathrm{i}Pk_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\mathrm{i}Pk_z & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathrm{i}Pk_x & \mathrm{i}Pk_y & \mathrm{i}Pk_z \\ 0 & 0 & 0 & 0 & -\mathrm{i}Pk_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mathrm{i}Pk_y & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mathrm{i}Pk_z & 0 & 0 & 0 \end{bmatrix}, \tag{A-3}$$

$$\mathbf{H}_2(\underline{\mathbf{k}}) = \begin{bmatrix} A_c(k_x^2 + k_y^2 + k_z^2) & 0 & 0 & 0 \\ 0 & \mathbf{D}_{3\times3}(\underline{\mathbf{k}}) & 0 & 0 \\ 0 & 0 & A_c(k_x^2 + k_y^2 + k_z^2) & 0 \\ 0 & 0 & 0 & \mathbf{D}_{3\times3}(\underline{\mathbf{k}}) \end{bmatrix}, \tag{A-4}$$

$$\mathbf{D}_{3\times3}(\underline{\mathbf{k}}) = \begin{bmatrix} Lk_x^2 + M(k_y^2 + k_z^2) & Nk_xk_y & k_xN_+k_z + k_zN_-k_x \\ Nk_xk_y & Lk_y^2 + M(k_x^2 + k_z^2) & k_yN_+k_z + k_zN_-k_y \\ k_zN_+k_x + k_xN_-k_z & k_zN_+k_y + k_yN_-k_z & Lk_z^2 + M(k_x^2 + k_y^2) \end{bmatrix}, \tag{A-5}$$

and

$$\mathbf{H}_\varepsilon = \begin{bmatrix} a_s(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) & 0 & 0 & 0 \\ 0 & \mathbf{D}_\varepsilon & 0 & 0 \\ 0 & 0 & a_s(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}) & 0 \\ 0 & 0 & 0 & \mathbf{D}_\varepsilon \end{bmatrix}, \tag{A-6}$$

$$\mathbf{D}_\varepsilon = \begin{bmatrix} l_\varepsilon\varepsilon_{xx} + m_\varepsilon(\varepsilon_{yy} + \varepsilon_{zz}) & n_\varepsilon\varepsilon_{xy} & n_\varepsilon\varepsilon_{xz} \\ n_\varepsilon\varepsilon_{yx} & l_\varepsilon\varepsilon_{xx} + m_\varepsilon(\varepsilon_{yy} + \varepsilon_{zz}) & n_\varepsilon\varepsilon_{yz} \\ n_\varepsilon\varepsilon_{zx} & n_\varepsilon\varepsilon_{zy} & l_\varepsilon\varepsilon_{xx} + m_\varepsilon(\varepsilon_{yy} + \varepsilon_{zz}) \end{bmatrix}, \tag{A-7}$$

is the strain induced deformation potential interaction. Under the effect of strain, the correction $\underline{\mathbf{k}} \to (1 - \varepsilon) \cdot \underline{\mathbf{k}}$ should also be applied. Here, E_g is the fundamental gap of the unstrained material, Δ is the spin-orbit splitting and P is the interband momentum matrix element, given by

$$P = \sqrt{\frac{\hbar^2}{2m_0} E_p}. \tag{A-8}$$

The renormalized parameters A_c, L, M, and N are obtained from the conduction band effective mass m_c^* and the modified Luttinger parameters γ_1, γ_2 and γ_3; in which the remote contribution of the conduction band has been subtracted [116, 91]

$$A_c = \frac{\hbar^2}{2m_c^*} - \frac{2P^2}{3E_g} - \frac{P^2}{3(E_g + \Delta)} \tag{A-9}$$

$$L = -\frac{\hbar^2}{2m_0}(\gamma_1 + 4\gamma_2) \tag{A-10}$$

$$M = -\frac{\hbar^2}{2m_0}(\gamma_1 - 2\gamma_2) \tag{A-11}$$

$$N = -\frac{\hbar^2}{2m_0}6\gamma_3. \tag{A-12}$$

According to operator ordering, the Kane parameter N is split into two asymmetric parts $N = N_+ + N_-$, with $N_- = M - \hbar^2/(2m_0)$. The splitting is shown only for the matrix elements of (A-5) that are linear in k_z, assuming z as the only symmetry-broken direction. Moving to the atomic-like basis set defined in [117, 118], and adopting the axial approximation, the bulk Hamiltonian block-diagonalizes with respect to the spin components (*i.e.*, the 8×8 Hamiltonian decouples in two 4×4 blocks), and the energy dispersion becomes isotropic in the transverse wavevector. The band parameters for some of the most important Sb-based alloys are shown in Tab. **A-1**.

Table A-1: Band parameters for AlSb, InAs, GaSb and InSb, taken from [119].

Parameters	AlSb	InAs	GaSb	InSb
E_g	2.386	0.42	0.81	0.235
m_c/m_0	0.13	0.0224	0.0412	0.0135
γ_1	4.15	19.4	11.84	32.4
γ_2	1.28	8.545	4.25	13.3
γ_3	1.75	9.17	5.01	15.15
Δ	0.65	0.38	0.81	0.81
E_p	18.7	21.5	22.4	23.3

A.2 Wurtzite model

The eight-band $\mathbf{k} \cdot \mathbf{p}$ bulk Hamiltonian for (0001)-oriented wurtzite semiconductors [36]

$$\mathbf{H}^{k \cdot p}_{\text{Bult}} = \begin{bmatrix} E_c & 0 & -k_+P_2 & k_-P_2 & k_zP_1 & 0 & 0 & 0 \\ 0 & E_c & 0 & 0 & 0 & k_-P_2 & -k_+P_2 & k_zP_1 \\ -k_-P_2 & 0 & F & -(K^\dagger)^* & -(H_+^\dagger)^* & 0 & 0 & 0 \\ k_+P_2 & 0 & -K^* & G & H_-^\dagger & 0 & 0 & \sqrt{2}\Delta_3 \\ k_zP_1 & 0 & -H_+^* & H_- & \Lambda & 0 & \sqrt{2}\Delta_3 & 0 \\ 0 & k_+P_2 & 0 & 0 & 0 & F^* & -K^* & H_-^\dagger \\ 0 & -k_-P_2 & 0 & 0 & \sqrt{2}\Delta_3 & -(K^\dagger)^* & G^* & -(H_+^\dagger)^* \\ 0 & k_zP_1 & 0 & \sqrt{2}\Delta_3 & 0 & H_- & -H_+^* & \Lambda \end{bmatrix}, \tag{A-13}$$

written in the zone-center basis

$$|1\rangle = |\mathrm{i}S\uparrow\rangle, \quad |2\rangle = |\mathrm{i}S\downarrow\rangle,$$
$$|3\rangle = \left|-\frac{1}{\sqrt{2}}(X+\mathrm{i}Y)\uparrow\right\rangle, \quad |4\rangle = \left|\frac{1}{\sqrt{2}}(X-\mathrm{i}Y)\uparrow\right\rangle, \quad |5\rangle = |Z\uparrow\rangle,$$
$$|6\rangle = \left|\frac{1}{\sqrt{2}}(X-\mathrm{i}Y)\downarrow\right\rangle, \quad |7\rangle = \left|-\frac{1}{\sqrt{2}}(X+\mathrm{i}Y)\downarrow\right\rangle, \quad |8\rangle = |Z\downarrow\rangle,$$

with

$$k_\pm = \frac{1}{\sqrt{2}}(k_x \pm \mathrm{i}k_y), \tag{A-14}$$

$$E_c = E_g + \Delta_1 + \Delta_2 + \frac{\hbar^2}{2m_0}[k_x A_t k_x + k_y A_t k_y + k_z A_z k_z], \tag{A-15}$$

$$F = \Delta_1 + \Delta_2 + \Lambda + \Theta + \frac{\hbar^2}{2m_0}[-\mathrm{i}k_y(A_5^+ - A_5^-)k_x + \mathrm{i}k_x(A_5^+ - A_5^-)k_y], \tag{A-16}$$

$$G = \Delta_1 - \Delta_2 + \Lambda + \Theta + \frac{\hbar^2}{2m_0}[-\mathrm{i}k_x(A_5^+ - A_5^-)k_y + \mathrm{i}k_y(A_5^+ - A_5^-)k_x], \tag{A-17}$$

$$\Lambda = \frac{\hbar^2}{2m_0}[k_z A_1 k_z + k_x A_2 k_x + k_y A_2 k_y] + \Lambda_\varepsilon, \tag{A-18}$$

$$\Lambda_\varepsilon = D_1\varepsilon_{zz} + D_2(\varepsilon_{xx} + \varepsilon_{xx}), \tag{A-19}$$

$$\Theta = \frac{\hbar^2}{2m_0}[k_z A_3 k_z + k_x A_4 k_x + k_y A_4 k_y] + \Theta_\varepsilon, \tag{A-20}$$

$$\Theta_\varepsilon = D_3\varepsilon_{zz} + D_4(\varepsilon_{xx} + \varepsilon_{xx}), \tag{A-21}$$

$$K = \frac{\hbar^2}{2m_0}[k_x A_5 k_x - \mathrm{i}(k_x A_5 k_y + k_x A_5 k_y) - k_y A_5 k_y] + D_5(\varepsilon_{xx} - 2\mathrm{i}\varepsilon_{xy} - \varepsilon_{yy}) \tag{A-22}$$

$$H_\pm = \frac{\hbar^2}{2m_0}[k_z A_6^+(k_x - \mathrm{i}k_y) + (k_x - \mathrm{i}k_y)A_6^- k_z] \pm \mathrm{i}A_7(k_x - \mathrm{i}k_y) + D_6(\varepsilon_{zx} - \mathrm{i}\varepsilon_{yz}). \tag{A-23}$$

Here $\Delta_1 = \Delta_{\mathrm{cr}}$ is the crystal field splitting energy, $\Delta_2 = \Delta_3 = \Delta_{\mathrm{SO}}/3$ with Δ_{SO} the spin-orbit splitting energy, $D_1, ..., D_6$ are deformation potentials accounting for strain effects, and $A_1, ..., A_7$ are the band structure parameters. The notation introduced by Veprek [120] has been used to indicate operator ordering, needed for numerical stability. By using the basis transformation $\mathbf{H}' = \mathbf{T}^*\mathbf{H}\mathbf{T}^T$, with

$$\mathbf{T} = \begin{bmatrix} \mathrm{i}\beta^* & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha^* & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 & \beta^* & 0 \\ 0 & 0 & 0 & 0 & \beta^* & 0 & 0 & \beta \\ \mathrm{i}\beta^* & -\alpha & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha^* & 0 & 0 & -\alpha & 0 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 & -\beta^* & 0 \\ 0 & 0 & 0 & 0 & -\beta^* & 0 & 0 & \beta \end{bmatrix}, \tag{A-24}$$

where $\alpha = e^{\mathrm{i}(3\pi/4+3\phi/2)}/\sqrt{2}$, $\beta = e^{\mathrm{i}(\pi/4+\phi/2)}/\sqrt{2}$ and $\phi = \mathrm{atan}(k_y/k_x)$, the bulk 8×8 Hamiltonian (A-13) is block diagonalized with respect to the spin components into two (real) 4×4 blocks, which simplifies the numerical evaluation of the Green's functions. The band structure parameters for some of the main nitride compounds are shown in Tab. **A-2**.

Table A-2: Band parameters for wurtzite GaN, AlN and InN, taken from [121].

Parameters	GaN	AlN	InN
E_g (eV)	3.44	6.16	0.69
m_e^z/m_0	0.186	0.322	0.065
m_e^t/m_0	0.209	0.329	0.068
A_1	-5.947	-3.991	-15.803
A_2	-0.528	-0.311	-0.497
A_3	5.414	3.671	15.251
A_4	-2.512	-1.147	-7.151
A_5	-2.510	-1.329	-7.060
A_6	-3.202	-1.952	-10.078
A_7 (eVÅ)	0.046	0.026	0.175
Δ_{cr} (meV)	-295	10	66
Δ_{so} (meV)	19	17	5

Appendix B

Carrier interaction with photons and phonons

In the following sections, the derivation of the self-energy for the interaction of electrons with bosons (photons and phonons) is presented, based on the work in [27, 46]. The procedure followed is similar to the one presented section 2.7, in which the interaction Hamiltonian is inserted into the perturbation expansion of the Green's function (2-133), and then the irreducible diagrams of the self-energy (2-148) are identified from the resulting expansion. We will restrict our approximation of the self-energy to the Hartree-Fock diagrams.

B.1 Interaction Hamiltonian and perturbation expansion

The interaction potential accounting for electron coupling with photons and phonons take the general form

$$V_{e-\gamma}(\mathbf{r},t) = \sum_{\mathbf{q}} M(\mathbf{q}) e^{\mathbf{i}\mathbf{q}\cdot\mathbf{r}} \left(\hat{b}_{\mathbf{q}}(t) + \hat{b}^{\dagger}_{-\mathbf{q}}(t) \right), \tag{B-1}$$

where $\hat{b}^{\dagger}_{\mathbf{q}}$ and $\hat{b}_{\mathbf{q}}$ are the bosonic creation and annihilation operators, respectively, and $\mathbf{M}(\mathbf{q})$ is a function of the 3D boson wavevector $\mathbf{q} = (\mathbf{q}_{||}, q_z)$ that depends on the particular type of interaction. In second quantization, the interaction Hamiltonian which will be introduced in the perturbation expansion, is given by

$$\hat{H}^{i}(t) = \int d\mathbf{r} \hat{\Psi}^{\dagger}(\mathbf{r},t) V_{e-\gamma}(\mathbf{r},t) \hat{\Psi}(\mathbf{r},t) \tag{B-2}$$

where the fermionic field operators are assumed to be expressed in the interaction picture ($\hat{\Psi} \equiv \hat{\Psi}_I$), and we have omitted the subscript for simplicity. Inserting the expansions (2-65a)

and (2-65b) of the fermionic field operators in terms of the discrete basis[1] $\{\phi_{\nu,\mathbf{k}}(\mathbf{r})\}$, with the basis functions given by (2-77), the interaction Hamiltonian (B-2) gives

$$\begin{aligned}\hat{H}^i(t) &= \sum_{\nu,\mu}\sum_{\mathbf{k},\mathbf{k}'}\int d\mathbf{r}\,\phi^*_{\nu,\mathbf{k}}(\mathbf{r})\hat{c}^\dagger_{\nu,\mathbf{k}}(t)\left[\sum_{\mathbf{q}} M(\mathbf{q})e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}\left(\hat{b}_{\mathbf{q}}(t)+\hat{b}^\dagger_{-\mathbf{q}}(t)\right)\right]\phi_{\mu,\mathbf{k}'}(\mathbf{r})\hat{c}_{\mu,\mathbf{k}'}(t) \\ &= \sum_{\nu,\mu}\sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} M_{\nu\mu}(\mathbf{k},\mathbf{k}',\mathbf{q})\hat{c}^\dagger_{\nu,\mathbf{k}}(t)\hat{c}_{\mu,\mathbf{k}'}(t)\left[\hat{b}_{\mathbf{q}}(t)+\hat{b}^\dagger_{-\mathbf{q}}(t)\right], \end{aligned} \tag{B-3}$$

where we have introduced the interaction matrix element

$$M_{\nu\mu}(\mathbf{k},\mathbf{k}',\mathbf{q}) = \int d\mathbf{r}\,\phi^*_{\nu,\mathbf{k}}(\mathbf{r})e^{\mathrm{i}\mathbf{q}\cdot\mathbf{r}}M(\mathbf{q})\phi_{\mu,\mathbf{k}'}(\mathbf{r}). \tag{B-4}$$

The interaction Hamiltonian (B-3) should be inserted in (2-133) and the exponential term expanded as a power series of the perturbation, providing a perturbative expansion of the Green's function for interactions of the electron-phonon/photon type. Since terms arising from the first-order perturbation (2-135b) (and, in general, any odd-order term) are proportional to either $\hat{b}$ or $\hat{b}^\dagger$, their respective average over the equilibrium ensemble vanish as it is proportional to $\langle\hat{b}\rangle_0 = \langle\hat{b}^\dagger\rangle_0 = 0$. The second-order contribution of the expansion reads

$$G^{(2)}(\mathbf{r},t;\mathbf{r}',t') = \left(\frac{-i}{\hbar}\right)^3\frac{1}{2}\langle\hat{T}_C\left\{\int_C ds\int_C ds' H^i(s)H^i(s')\hat{\Psi}(\mathbf{r},t)\hat{\Psi}^\dagger(\mathbf{r}',t')\right\}\rangle, \tag{B-5}$$

where the subscript for an equilibrium ensemble average has been dropped for simplicity ($\langle\ldots\rangle \equiv \langle\ldots\rangle_0$). Expanding (B-5) in terms of the fermion operators $\hat{c}$ and $\hat{c}^\dagger$ we obtain

$$G^{(2)}(\mathbf{r},t;\mathbf{r}',t') = \sum_{\nu\mu}\sum_{\mathbf{k}}\phi_{\nu,\mathbf{k}}(\mathbf{r})\left(\frac{-i}{\hbar}\right)^3\frac{1}{2}\langle\hat{T}_C\left\{\int_C ds\int_C ds' H^i(s)H^i(s')\hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\mu,\mathbf{k}}(t')\right\}\rangle\phi^*_{\mu,\mathbf{k}}(\mathbf{r}'), \tag{B-6}$$

from which a direct comparison with (2-80) allows to identify the contravariant representation of the Green's function

$$G^{(2)}_{\nu\mu}(\mathbf{k};t,t') = \left(\frac{-i}{\hbar}\right)^3\frac{1}{2}\langle\hat{T}_C\left\{\int_C ds\int_C ds' H^i(s)H^i(s')\hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\mu,\mathbf{k}}(t')\right\}\rangle. \tag{B-7}$$

Further substituting the interaction Hamiltonian (B-3) in (B-7), we obtain

$$\begin{aligned}G^{(2)}_{\nu\mu}(\mathbf{k};t,t') = \left(\frac{-i}{\hbar}\right)^3\frac{1}{2}\langle\hat{T}_C\Big\{\int_C ds\int_C ds' &\sum_{\nu_1,\nu_2}\sum_{\mu_1,\mu_2}\sum_{\mathbf{k}_1,\mathbf{k}_3,\mathbf{q}_1}\sum_{\mathbf{k}_2,\mathbf{k}_4,\mathbf{q}_2} M_{\nu_1\mu_1}(\mathbf{k}_1,\mathbf{k}_3,\mathbf{q}_1) \\ &\times M_{\nu_2\mu_2}(\mathbf{k}_2,\mathbf{k}_4,\mathbf{q}_2)\left[\hat{b}_{\mathbf{q}_1}(s)+\hat{b}^\dagger_{-\mathbf{q}_1}(s)\right]\left[\hat{b}_{\mathbf{q}_2}(s')+\hat{b}^\dagger_{-\mathbf{q}_2}(s')\right] \\ &\times \hat{c}^\dagger_{\nu_1,\mathbf{k}_1}(s)\hat{c}_{\mu_1,\mathbf{k}_3}(s)\hat{c}^\dagger_{\nu_2,\mathbf{k}_2}(s')\hat{c}_{\mu_2,\mathbf{k}_4}(s')\hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\mu,\mathbf{k}}(t')\Big\}\rangle. \end{aligned} \tag{B-8}$$

[1] As a small reminder for the reader, the subindex $\nu = (i, m)$ combines the space i and band m indices, and $\mathbf{k}$ denotes the transversal electron crystal momentum.

Now, we can apply *Wick-Matsubara* theorem [18] to factor the electron operators product. The boson operator product become [27]

$$\begin{aligned}
&\langle \hat{T}_C \left\{ \left[\hat{b}_{\mathbf{q}_1}(s) + \hat{b}^\dagger_{-\mathbf{q}_1}(s)\right] \left[\hat{b}_{\mathbf{q}_2}(s') + \hat{b}^\dagger_{-\mathbf{q}_2}(s')\right] \right\}\rangle \\
&= \langle \hat{T}_C \left\{ \hat{b}_{\mathbf{q}_1}(s)\hat{b}^\dagger_{-\mathbf{q}_2}(s') \right\}\rangle + \langle \hat{T}_C \left\{ \hat{b}^\dagger_{-\mathbf{q}_1}(s)\hat{b}_{\mathbf{q}_2}(s') \right\}\rangle \\
&= \mathrm{i}\hbar D_0(\mathbf{q}_1; s, s')\delta_{\mathbf{q}_1,-\mathbf{q}_2},
\end{aligned} \tag{B-9}$$

where products involving two creation or two annihilation operators vanish from the sum, and we have introduced the free boson Green's function defined as [122]

$$D_0(\mathbf{q}; t, t') = -\frac{i}{\hbar}\left[\langle \hat{T}_C \left\{ \hat{b}_{\mathbf{q}}(t)\hat{b}^\dagger_{\mathbf{q}}(t') \right\}\rangle + \langle \hat{T}_C \left\{ \hat{b}^\dagger_{-\mathbf{q}}(t)\hat{b}_{-\mathbf{q}}(t') \right\}\rangle\right]. \tag{B-10}$$

In (B-9) we have used that the boson wavevector should remain conserved when evaluating the equilibrium ensemble average due to the absence of interactions, *i.e.*,

$$\langle \hat{T}_C \left\{ \hat{b}_{\mathbf{q}_1}(s)\hat{b}^\dagger_{-\mathbf{q}_2}(s') \right\}\rangle \propto \delta_{\mathbf{q}_1,-\mathbf{q}_2}, \tag{B-11}$$

with $\delta_{\mathbf{q}_1,-\mathbf{q}_2}$ the Kronecker delta function. Analogous to the procedure followed in Sec. 2.7, the product of the three pairs of creation and annihilation electron operators give rise to six terms: two correspond to disconnected diagrams (this terms can be factored and cancel out with the diagrams in the denominator [21]), while the remaining four correspond to equivalent factorizations of the Hartree and Fock diagrams, cancelling out the factor one half. Therefore, the factorization of the electron operators give the result [27]

$$\begin{aligned}
&\frac{1}{2}\langle \hat{T}_C \left\{ \hat{c}^\dagger_{\nu_1,\mathbf{k}_1}(s)\hat{c}_{\mu_1,\mathbf{k}_3}(s)\hat{c}^\dagger_{\nu_2,\mathbf{k}_2}(s')\hat{c}_{\mu_2,\mathbf{k}_4}(s')\hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\mu,\mathbf{k}}(t') \right\}\rangle \\
&= \langle \hat{T}_C \left\{ \hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\nu_1,\mathbf{k}_1}(s) \right\}\rangle \langle \hat{T}_C \left\{ \hat{c}_{\mu_1,\mathbf{k}_3}(s)\hat{c}^\dagger_{\nu_2,\mathbf{k}_2}(s') \right\}\rangle \langle \hat{T}_C \left\{ \hat{c}_{\mu_2,\mathbf{k}_4}(s')\hat{c}^\dagger_{\mu,\mathbf{k}}(t') \right\}\rangle \\
&- \langle \hat{T}_C \left\{ \hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\nu_1,\mathbf{k}_1}(s) \right\}\rangle \langle \hat{T}_C \left\{ \hat{c}_{\mu_2,\mathbf{k}_4}(s')\hat{c}^\dagger_{\nu_2,\mathbf{k}_2}(s') \right\}\rangle \langle \hat{T}_C \left\{ \hat{c}_{\mu_1,\mathbf{k}_3}(s)\hat{c}^\dagger_{\mu,\mathbf{k}}(t') \right\}\rangle \\
&= \mathrm{i}\hbar G_{0,\nu\nu_1}(\mathbf{k}; t, s)\delta_{\mathbf{k},\mathbf{k}_1}\, i\hbar G_{0,\mu_1\nu_2}(\mathbf{k}_2; s, s')\delta_{\mathbf{k}_3,\mathbf{k}_2}\, i\hbar G_{0,\mu_2\mu}(\mathbf{k}; s', t')\delta_{\mathbf{k}_4,\mathbf{k}} \\
&- i\hbar G_{0,\nu\nu_1}(\mathbf{k}; t, s)\delta_{\mathbf{k},\mathbf{k}_1}\, i\hbar G_{0,\mu_2\nu_2}(\mathbf{k}_2; s', s')\delta_{\mathbf{k}_4,\mathbf{k}_2}\, i\hbar G_{0,\mu_1\mu}(\mathbf{k}; s, t')\delta_{\mathbf{k}_3,\mathbf{k}},
\end{aligned} \tag{B-12}$$

where we have introduced the noninteracting Green's function (2-12) in its contravariant representation

$$G_{0,\nu\mu}(\mathbf{k}; t, t') = -\frac{i}{\hbar}\langle \hat{T}_C \left\{ \hat{c}_{\nu,\mathbf{k}}(t)\hat{c}^\dagger_{\mu,\mathbf{k}}(t') \right\}\rangle. \tag{B-13}$$

Similar to the case of the boson operator, the conservation of the electron wavevector in Eq. (B-12) is expressed through the Kronecker delta functions. Inserting Eqs. (B-9) and (B-12)

in (B-8) (with $\mathbf{q}_1 = -\mathbf{q}_2 = \mathbf{q}$ and $\mathbf{k}_2 = \mathbf{k}'$), yields

$$\begin{aligned} G^{(2)}_{\nu,\mu}(\mathbf{k};t,t') = \int_C ds \int_C ds' \sum_{\nu_1,\nu_2} \sum_{\mu_1,\mu_2} \sum_{\mathbf{k}',\mathbf{q}} & \mathrm{i}\hbar D_0(\mathbf{q};s,s') \\ & \times [M_{\nu_1\mu_1}(\mathbf{k},\mathbf{k}',\mathbf{q}) M_{\nu_2\mu_2}(\mathbf{k}',\mathbf{k},-\mathbf{q}) \\ & \times G_{0,\nu\nu_1}(\mathbf{k};t,s)\, G_{0,\mu_1\nu_2}(\mathbf{k}';s,s')\, G_{0,\mu_2\mu}(\mathbf{k};s',t') \\ & - M_{\nu_1\mu_1}(\mathbf{k},\mathbf{k},\mathbf{q}) M_{\nu_2\mu_2}(\mathbf{k}',\mathbf{k}',-\mathbf{q}) \\ & \times G_{0,\nu\nu_1}(\mathbf{k};t,s)\, G_{0,\mu_2\nu_2}(\mathbf{k}';s',s')\, G_{0,\mu_1\mu}(\mathbf{k};s,t')]. \end{aligned} \tag{B-14}$$

The full Green's function up to second order interaction is then given by

$$\begin{aligned} G(\mathbf{r},t;\mathbf{r}',t') &\approx G_0(\mathbf{r},t;\mathbf{r}',t') + G^{(2)}(\mathbf{r},t;\mathbf{r}',t') \\ &= \sum_{\nu\mu} \sum_{\mathbf{k}} \phi_{\nu,\mathbf{k}}(\mathbf{r}) G_{0,\nu\mu}(\mathbf{k};t,t') \phi^*_{\mu,\mathbf{k}}(\mathbf{r}') + \sum_{\nu\mu} \sum_{\mathbf{k}} \phi_{\nu,\mathbf{k}}(\mathbf{r}) G^{(2)}_{\nu\mu}(\mathbf{k};t,t') \phi^*_{\mu,\mathbf{k}}(\mathbf{r}') \\ &= \sum_{\nu\mu} \sum_{\mathbf{k}} \phi_{\nu,\mathbf{k}}(\mathbf{r}) \left[G_{0,\nu\mu}(\mathbf{k};t,t') + G^{(2)}_{\nu\mu}(\mathbf{k};t,t') \right] \phi^*_{\mu,\mathbf{k}}(\mathbf{r}'), \end{aligned} \tag{B-15}$$

with its contravariant representation in the discrete basis $\{\phi_{\nu,\mathbf{k}}(\mathbf{r})\}$ given by

$$G_{\nu\mu}(\mathbf{k};t,t') \approx G_{0,\nu\mu}(\mathbf{k};t,t') + G^{(2)}_{\nu\mu}(\mathbf{k};t,t'). \tag{B-16}$$

B.2 Derivation of the self-energy

Equation (B-14) for the second-order contribution of the perturbation expansion of the Green's function can be written, after reorganizing and factoring terms, in the following way

$$\begin{aligned} G^{(2)}_{\nu\mu}(\mathbf{k};t,t') = \int_C ds \int_C ds' & \sum_{\nu_1,\mu_1} G_{0,\nu\nu_1}(\mathbf{k};t,s) \\ & \times \Bigg[\mathrm{i}\hbar \sum_{\mathbf{k}',\mathbf{q}} D_0(\mathbf{q};s,s') \sum_{\mu_2,\nu_2} M_{\nu_1\mu_2}(\mathbf{k},\mathbf{k}',\mathbf{q})\, G_{0,\mu_2\nu_2}(\mathbf{k}';s,s') M_{\nu_2\mu_1}(\mathbf{k}',\mathbf{k},-\mathbf{q}) \\ & - \mathrm{i}\hbar\, \delta_C(s,s') \sum_{\mathbf{k}',\mathbf{q}} M_{\nu_1\mu_1}(\mathbf{k},\mathbf{k},\mathbf{q}) \int_C d\tau\, D_0(\mathbf{q};s,\tau) \sum_{\mu_2,\nu_2} G_{0,\mu_2\nu_2}(\mathbf{k}';\tau,\tau) M_{\nu_2\mu_2}(\mathbf{k}',\mathbf{k}',-\mathbf{q}) \Bigg] \\ & \times G_{0,\mu_1\mu}(\mathbf{k};s',t'), \end{aligned} \tag{B-17}$$

or equivalently, in full matrix notation

$$\mathbf{G}^{(2)}(\mathbf{k};t,t') = \int_C ds \int_C ds'\, \mathbf{G}_0(\mathbf{k};t,s) \boldsymbol{\Sigma}(\mathbf{k};s,s')\, \mathbf{G}_0(\mathbf{k};s',t'). \tag{B-18}$$

A direct comparison of Eqs. (B-16) and (B-18) with (2-142) allows to identify the Hartree-Fock self-energy, which reads in full matrix notation

$$\begin{aligned}\mathbf{\Sigma}(\mathbf{k};t,t') = & \underbrace{-\mathrm{i}\hbar\,\delta_C(t,t')\sum_{\mathbf{k}',\mathbf{q}}\mathbf{M}(\mathbf{k},\mathbf{k},\mathbf{q})\int_C ds\,D_0(\mathbf{q};t,s)\,\mathrm{Tr}[\mathbf{G}_0(\mathbf{k}';s,s)\,\mathbf{M}(\mathbf{k}',\mathbf{k}',-\mathbf{q})]}_{\mathbf{\Sigma}_H(\mathbf{k};t,t')} \\ & \underbrace{+\,\mathrm{i}\hbar\sum_{\mathbf{k}',\mathbf{q}}D_0(\mathbf{q};t,t')\mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\mathbf{G}_0(\mathbf{k}';t,t')\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q})}_{\mathbf{\Sigma}_F(\mathbf{k},t,t')}\,.\end{aligned} \tag{B-19}$$

The two terms composing the self-energy in (B-19) correspond to the lowest order diagrams, *i.e.*, the Hartree self-energy $\mathbf{\Sigma}_H$ (a bubble diagram can be recognized from the free Green's function initiating and ending at the same point) and the Fock self-energy $\mathbf{\Sigma}_F$ from equation (2-143). The Hartree self-energy can be demonstrated to be identically zero in bulk systems [27], and for the case of heterostructures it has been shown to be important for some phenomena, *e.g.*, the polaron shift in resonant tunnelling diodes operating at regimes where the resonant level is completely occupied [122]. Similar to the procedure followed in section 2.7, we use the *self-consistent Born approximation* (SCBA) by replacing the noninteracting Green's function $\mathbf{G}_0$ with the exact Green's function $\mathbf{G}$, leading for carrier-boson Hartree-Fock self-energy

$$\mathbf{\Sigma}_H(\mathbf{k};t,t') = -\mathrm{i}\hbar\,\delta_C(t,t')\sum_{\mathbf{k}',\mathbf{q}}\mathbf{M}(\mathbf{k},\mathbf{k},\mathbf{q})\int_C ds\,D_0(\mathbf{q};t,s)\,\mathrm{Tr}[\mathbf{G}(\mathbf{k}';s,s)\,\mathbf{M}(\mathbf{k}',\mathbf{k}',-\mathbf{q})] \tag{B-20}$$

$$\mathbf{\Sigma}_F(\mathbf{k};t,t') = \mathrm{i}\hbar\sum_{\mathbf{k}',\mathbf{q}}D_0(\mathbf{q};t,t')\mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\mathbf{G}(\mathbf{k}';t,t')\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}), \tag{B-21}$$

The Langreth rules[2] are applied to obtain the real-time components of the self-energy (lesser, greater and retarded), yielding

$$\mathbf{\Sigma}_H^{\lessgtr}(\mathbf{k};t,t') = 0 \tag{B-22a}$$

$$\mathbf{\Sigma}_H^{R}(\mathbf{k};t,t') = -\mathrm{i}\hbar\,\delta(t-t')\sum_{\mathbf{k}',\mathbf{q}}\mathbf{M}(\mathbf{k},\mathbf{k},\mathbf{q})\int_{-\infty}^{\infty} ds\,D_0^R(\mathbf{q};t,s)\,\mathrm{Tr}[\mathbf{G}^<(\mathbf{k}';s,s)\,\mathbf{M}(\mathbf{k}',\mathbf{k}',-\mathbf{q})], \tag{B-22b}$$

[2]A summary of the Langreth rules can be found in [20], Ch. 4.3.

for the Hartree self-energy, and

$$\boldsymbol{\Sigma}_F^{\lessgtr}(\mathbf{k};t,t') = \mathrm{i}\hbar \sum_{\mathbf{k}',\mathbf{q}} D_0^{\lessgtr}(\mathbf{q};t,t')\mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\mathbf{G}^{\lessgtr}(\mathbf{k}';t,t')\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}) \tag{B-23a}$$

$$\boldsymbol{\Sigma}_F^{R}(\mathbf{k};t,t') = \mathrm{i}\hbar \sum_{\mathbf{k}',\mathbf{q}} \mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\left[D_0^{R}(\mathbf{q};t,t')\,\mathbf{G}^{R}(\mathbf{k}';t,t') + D_0^{R}(\mathbf{q};t,t')\,\mathbf{G}^{<}(\mathbf{k}';t,t') + D_0^{<}(\mathbf{q};t,t')\,\mathbf{G}^{R}(\mathbf{k}';t,t')\right]\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}), \tag{B-23b}$$

for the Fock self-energy. Note that the lesser/greater components of the Hartree self-energy $\boldsymbol{\Sigma}_H^{\lessgtr}$ vanish because the times are by definition in different branches of the time contour and the Dirac delta is always zero. The steady-state expressions of the self-energies are obtained by Fourier transforming (B-22b), (B-23a) and (B-23b) with respect to the time difference $\tau = t - t'$, obtaining

$$\boldsymbol{\Sigma}_H^{R}(\mathbf{k}) = -\mathrm{i} \sum_{\mathbf{k}',\mathbf{q}} \mathbf{M}(\mathbf{k},\mathbf{k},\mathbf{q})\,D_0^{R}(\mathbf{q},E=0) \int \frac{dE'}{2\pi} \mathrm{Tr}[\mathbf{G}^{<}(\mathbf{k}',E')\,\mathbf{M}(\mathbf{k}',\mathbf{k}',-\mathbf{q})] \tag{B-24}$$

$$\boldsymbol{\Sigma}_F^{\lessgtr}(\mathbf{k},E) = \mathrm{i} \sum_{\mathbf{k}',\mathbf{q}} \int \frac{dE'}{2\pi}\, D_0^{\lessgtr}(\mathbf{q},E')\mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\mathbf{G}^{\lessgtr}(\mathbf{k}',E-E')\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}) \tag{B-25}$$

$$\boldsymbol{\Sigma}_F^{R}(\mathbf{k},E) = \mathrm{i} \sum_{\mathbf{k}',\mathbf{q}} \int \frac{dE'}{2\pi}\, \mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q})\,\left[D_0^{R}(\mathbf{q},E')\,\mathbf{G}^{R}(\mathbf{k}',E-E') + D_0^{R}(\mathbf{q},E')\,\mathbf{G}^{<}(\mathbf{k}',E-E') + D_0^{<}(\mathbf{q},E')\,\mathbf{G}^{R}(\mathbf{k}',E-E')\right]\mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}), \tag{B-26}$$

where the Hartree self-energy is energy independent.

B.2.1 Renormalization of the boson Green's function

The self-energies (B-20) and (B-21) consider an equilibrium population of bosons through the free boson propagator $D_0(\mathbf{q};t,t')$, *i.e.*, bosons are assumed to remain unperturbed by the interaction with the electron population. In reality, boson interaction with electrons should lead to a renormalization of the bosonic Green's function, and therefore the free boson propagator should be replaced by the full boson Green's function $D(\mathbf{q};t,t')$ accounting for electron-boson interaction. Such a substitution is equivalent to the procedure followed in the SCBA, where the free electron propagator is replaced by the full propagator inside the self-energies. Similar to the electron Green's function, the boson Green's function satisfies the Dyson's equation, which can be expressed diagrammatically as[3]

(B-27)

[3] The reader is referred to [23], Sec. 4.2.5, where the author derives the boson Green's function, including electron-boson interaction up to second order perturbation.

In Eq. (B-27), we have used the notation from table **2-1**. The double wiggly line represents the full boson Green's function. The boson self-energy $\Pi(\mathbf{q};t,t')$, often called *polarizability* [21], reads in the SCBA

(B-28)

where we have included just the first two diagrams coming second and fourth-order perturbation (odd order perturbation terms vanish because they contain an odd number of boson creation/annihilation operators). Inserting (B-28) in (B-27), gives for the full boson propagator

(B-29)

whose first two terms on the right hand side, when inserted in Eq. (B-21), correspond to the Fock and direct collision terms shown in the diagrammatic expansion (2-148). Therefore, solving the Dyson's equation for the boson propagator is equivalent to introducing higher order diagrams in the self-energy expansion. In the simplest approximation, assuming an equilibrium the boson population, the boson operator is given by [27]

$$\hat{b}_{\mathbf{q}}(t) = \hat{b}_{\mathbf{q}} e^{-\mathrm{i}\omega_{\mathbf{q}} t}, \tag{B-30}$$

and the free boson lesser and greater Green's functions can be directly computed as

$$\begin{aligned} D_0^{\lessgtr}(\mathbf{q};t,t') &= -\frac{\mathrm{i}}{\hbar}\left[\langle \hat{T}_C\left\{\hat{b}_{\mathbf{q}}(t)\hat{b}^\dagger_{\mathbf{q}}(t')\right\}\rangle + \langle \hat{T}_C\left\{\hat{b}^\dagger_{-\mathbf{q}}(t)\hat{b}_{-\mathbf{q}}(t')\right\}\rangle\right] \\ &= -\frac{\mathrm{i}}{\hbar}\left[\langle \hat{b}_{\mathbf{q}}\hat{b}^\dagger_{\mathbf{q}}\rangle e^{\pm\mathrm{i}\omega_{\mathbf{q}}(t-t')} + \langle \hat{b}^\dagger_{\mathbf{q}}\hat{b}_{\mathbf{q}}\rangle e^{\mp\mathrm{i}\omega_{\mathbf{q}}(t-t')}\right] \\ &= -\frac{\mathrm{i}}{\hbar}\left[(n_{\mathbf{q}}+1)e^{\pm\mathrm{i}\omega_{\mathbf{q}}(t-t')} + n_{\mathbf{q}} e^{\mp\mathrm{i}\omega_{\mathbf{q}}(t-t')}\right], \end{aligned} \tag{B-31}$$

with $n_{\mathbf{q}} = \langle \hat{b}^\dagger_{\mathbf{q}}\hat{b}_{\mathbf{q}}\rangle$ the average phonon number with wavevector $\mathbf{q}$, which in equilibrium is given by the Bose-Einstein distribution. The retarded component is given by

$$\begin{aligned} D_0^R(\mathbf{q};t,t') &= \theta(t-t')\left[D_0^>(\mathbf{q};t,t') - D_0^<(\mathbf{q};t,t')\right] \\ &= -\frac{2}{\hbar}\theta(t-t')\sin(\omega_{\mathbf{q}}(t-t')). \end{aligned} \tag{B-32}$$

Fourier transforming with respect to the time difference gives the steady-state expressions for the boson Green's function

$$D_0^{\lessgtr}(\mathbf{q},E) = -2\pi\mathrm{i}\left[(n_{\mathbf{q}}+1)\,\delta(E\pm\hbar\omega_{\mathbf{q}}) + n_{\mathbf{q}}\,\delta(E\mp\hbar\omega_{\mathbf{q}})\right] \tag{B-33a}$$

$$D_0^{R,A}(\mathbf{q},E) = \frac{1}{E-\hbar\omega_{\mathbf{q}}\pm i\eta} - \frac{1}{E+\hbar\omega_{\mathbf{q}}\pm i\eta}, \tag{B-33b}$$

which can be further substituted into the self-energies (B-24), (B-25) and (B-26), giving the final expressions

$$\mathbf{\Sigma}_H^R(\mathbf{k}) = \mathrm{i} \sum_{\mathbf{k}',\mathbf{q}} \frac{2}{\hbar\omega_{\mathbf{q}}} \mathbf{M}(\mathbf{k},\mathbf{k},\mathbf{q}) \int \frac{dE'}{2\pi} \mathrm{Tr}[\mathbf{G}^<(\mathbf{k}',E')\,\mathbf{M}(\mathbf{k}',\mathbf{k}',-\mathbf{q})] \tag{B-34}$$

$$\mathbf{\Sigma}_F^{\lessgtr}(\mathbf{k},E) = \sum_{\mathbf{k}',\mathbf{q}} \mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q}) \left[(n_{\mathbf{q}}+1)\mathbf{G}^{\lessgtr}(\mathbf{k}',E \pm \hbar\omega_{\mathbf{q}}) + n_{\mathbf{q}}\mathbf{G}^{\lessgtr}(\mathbf{k}',E \mp \hbar\omega_{\mathbf{q}})\right] \mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}) \tag{B-35}$$

$$\begin{aligned}\mathbf{\Sigma}_F^R(\mathbf{k},E) = \sum_{\mathbf{k}',\mathbf{q}} \mathbf{M}(\mathbf{k},\mathbf{k}',\mathbf{q}) \Big[& \frac{1}{2}\big(\mathbf{G}^<(\mathbf{k}',E-\hbar\omega_{\mathbf{q}}) - \mathbf{G}^<(\mathbf{k}',E+\hbar\omega_{\mathbf{q}})\big) \\ & + \mathrm{i}\mathcal{P}\left\{ \int \frac{dE'}{2\pi} \left(\frac{\mathbf{G}^<(\mathbf{k}',E-E')}{E'-\hbar\omega_{\mathbf{q}}} - \frac{\mathbf{G}^<(\mathbf{k}',E-E')}{E'+\hbar\omega_{\mathbf{q}}} \right) \right\} \\ & + (n_{\mathbf{q}}+1)\mathbf{G}^R(\mathbf{k}',E+\hbar\omega_{\mathbf{q}}) + n_{\mathbf{q}}\mathbf{G}^R(\mathbf{k}',E-\hbar\omega_{\mathbf{q}}) \Big] \mathbf{M}(\mathbf{k}',\mathbf{k},-\mathbf{q}),\end{aligned} \tag{B-36}$$

where everything has been written in full matrix notation.

Bibliography

[1] G. Ghione, *Semiconductor Devices for High-Speed Optoelectronics*. Cambridge, U.K.: Cambridge University Press, 2009.

[2] J. Piprek, "Efficiency droop in nitride-based light-emitting diodes," *Phys. Status Solidi A*, vol. 207, pp. 2217–2225, Oct. 2010.

[3] G. Verzellesi, D. Saguatti, M. Meneghini, F. Bertazzi, M. Goano, G. Meneghesso, and E. Zanoni, "Efficiency droop in InGaN/GaN blue light-emitting diodes: Physical mechanisms and remedies," *J. Appl. Phys.*, vol. 114, p. 071101, Aug. 2013.

[4] C. Weisbuch, M. Piccardo, L. Martinelli, J. Iveland, J. Peretti, and J. S. Speck, "The efficiency challenge of nitride light-emitting diodes for lighting," *Phys. Status Solidi A*, vol. 212, no. 5, pp. 899–913, 2015.

[5] A. David, N. G. Young, C. A. Hurni, and M. D. Craven, "Quantum efficiency of III-nitride emitters: Evidence for defect-assisted nonradiative recombination and its effect on the green gap," *Phys. Rev. Applied*, vol. 11, p. 031001, 2019.

[6] A. C. Espenlaub, D. J. Myers, E. C. Young, S. Marcinkevicius, C. Weisbuch, and J. S. Speck, "Evidence of trap-assisted auger recombination in low radiative efficiency MBE-grown III-nitride LEDs," *J. Appl. Phys.*, vol. 16, p. 184502, 2019.

[7] A. David, N. G. Young, C. Lund, and M. D. Craven, "Review-the physics of recombinations in III-nitride emitters," *ECS J. Solid State Sci. Technol.*, vol. 9, p. 016021, 2020.

[8] J. Hader, J. V. Moloney, and S. W. Koch, "Density-activated defect recombination as a possible explanation for the efficiency droop in GaN-based diodes," *Appl. Phys. Lett.*, vol. 96, p. 221106, 2010.

[9] P. Vogl and T. Kubis, "The non-equilibrium Green's function method: an introduction," *J. Comp. Electron.*, vol. 9, no. 3, pp. 237–242, 2010.

[10] M. Kira and S. W. Koch, "Many-body correlations and excitonic effects in semiconductor spectroscopy," *Progress Quantum Electron.*, vol. 30, no. 5, pp. 155–296, 2006.

[11] L. V. Keldysh, "Real-time nonequilibrium Green's functions," in *Progress in Nonequilibrium Green's Functions* (M. Bonitz and D. Semkat, eds.), (New Jersey), pp. 4–17, World Scientific, 2003.

[12] M. Auf der Maur, B. Galler, I. Pietzonka, M. Strassburg, H. Lugauer, and A. Di Carlo, "Trap-assisted tunneling in InGaN/GaN single-quantum-well light-emitting diodes," *Appl. Phys. Lett.*, vol. 105, no. 13, p. 133504, 2014.

[13] M. Mandurrino, G. Verzellesi, M. Goano, M. Vallone, F. Bertazzi, G. Ghione, M. Meneghini, G. Meneghesso, and E. Zanoni, "Physics-based modeling and experimental implications of trap-assisted tunneling in InGaN/GaN light-emitting diodes," *Phys. Status Solidi A*, vol. 212, no. 5, pp. 947–953, 2015.

[14] M. Goano, F. Bertazzi, X. Zhou, M. Mandurrino, S. Dominici, M. Vallone, G. Ghione, A. Tibaldi, M. Calciati, P. Debernardi, F. Dolcini, F. Rossi, G. Verzellesi, M. Meneghini, N. Trivellin, C. De Santi, E. Zanoni, and E. Bellotti, "Challenges towards the simulation of GaN-based LEDs beyond the semiclassical framework," in *SPIE Photonics West, Physics and Simulation of Optoelectronic Devices XXIV* (B. Witzigmann, M. Osiński, and Y. Arakawa, eds.), vol. 9742, Proceedings of the SPIE, (San Francisco, CA), p. 974202, Feb. 2016.

[15] C. De Santi, M. Buffolo, N. Renso, A. Neviani, G. Meneghesso, E. Zanoni, and M. Meneghini, "Evidence for defect-assisted tunneling and recombination at extremely low current in InGaN/GaN-based LEDs," *Appl. Phys. Express*, vol. 12, no. 5, p. 052007, 2019.

[16] M. Meneghini, C. De Santi, A. Tibaldi, M. Vallone, F. Bertazzi, G. Meneghesso, E. Zanoni, and M. Goano, "Thermal droop in III-nitride based light-emitting diodes: Physical origin and perspectives," *J. Appl. Phys.*, vol. 127, p. 211102, 2020.

[17] A. Wacker, "Semiconductor superlattices: A model system for nonlinear transport," *Phys. Rep.*, vol. 357, p. 1, 2002.

[18] C. Jacoboni, *Theory of Electron Transport in Semiconductors. A Pathway from Elementary Physics to Nonequilibrium Green Functions.* Berlin: Springer-Verlag, 2010.

[19] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics. Green's Function Methods in Equilibrium and Nonequilibrium Problems.* New York: W. A. Benjamin, 1962.

[20] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors.* Berlin: Springer, 2nd ed., 2008.

[21] G. Stefanucci and R. van Leeuwen, *Nonequilibrium Many-Body Theory Of Quantum Systems. A Modern Introduction.* Cambridge, U.K.: Cambridge University Press, 2013.

[22] J. Rammer, *Quantum Field Theory of Non-equilibrium States.* Cambridge, U.K.: Cambridge University Press, 2007.

[23] U. Aeberhard, *A Microscopic Theory of Quantum Well Photovoltaics.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2008.

[24] P. Danielewicz, "Quantum theory of nonequilibrium processes. I," *Ann. Phys.*, vol. 152, pp. 239–304, 1984.

[25] L. V. Keldysh, "Diagram technique for nonequilibrium processes," *Sov. Phys. JETP*, vol. 20, no. 4, p. 1018, 1965.

[26] F. S. Khan, J. H. Davies, and J. W. Wilkins, "Quantum transport equations for high electric fields," *Phys. Rev. B*, vol. 36, no. 5, p. 2578, 1987.

[27] R. K. Lake and R. R. Pandey, "Non-equilibrium Green functions in electronic device modeling," *arXiv:cond-mat/0607219v1*, 2007.

[28] U. Aeberhard, "Quantum-kinetic theory of photocurrent generation via direct and phonon-mediated optical transitions," *Phys. Rev. B*, vol. 84, p. 035454, July 2011.

[29] J. Piprek, ed., *Handbook of Optoelectronic Device Modeling and Simulation.* Boca Raton, FL: CRC Press, 2017.

[30] R. G. Veprek, *Computational modeling of semiconductor nanostructures for optoelectronics.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2009.

[31] S. Steiger, *Modelling Nano-LEDs.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2009.

[32] W. Schäfer and M. Wegener, *Semiconductor Optics and Transport Phenomena.* Springer, 2002.

[33] S. L. Chuang, *Physics of Photonic Devices.* New Jersey: John Wiley & Sons, 2009.

[34] J. M. Luttinger and W. Kohn, "Motion of electrons and holes in perturbed periodic fields," *Phys. Rev.*, vol. 97, pp. 869–883, 1955.

[35] P.-F. Qiao, S. Mou, and S. L. Chuang, "Electronic band structures and optical properties of type-II superlattice photodetectors with interfacial effect," *Opt. Express*, vol. 20, no. 3, p. 2319, 2012.

[36] X. Zhou, F. Bertazzi, M. Goano, G. Ghione, and E. Bellotti, "Deriving $k \cdot p$ parameters from full-Brillouin-zone descriptions: A finite-element envelope function model for quantum-confined wurtzite nanostructures," *J. Appl. Phys.*, vol. 116, p. 033709, July 2014.

[37] L. C. Lew Yan Voon and L. R. Ram-Mohan, "Tight-binding representation of the optical matrix elements: Theory and applications," *Phys. Rev. B*, vol. 47, no. 23, pp. 15500–15508, 1993.

[38] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, pp. 7845–7869, Feb. 1997.

[39] R. C. Bowen, W. R. Frensley, G. Klimeck, and R. K. Lake, "Transmission resonances and zeros in multiband models," *Phys. Rev. B*, vol. 52, no. 4, p. 2754, 1995.

[40] E. Polizzi and S. Datta, "Multidimensional nanoscale device modeling: the finite element method applied to the non-equilibrium Green's function formalism," in *3rd IEEE Conference on Nanotechnology (IEEE-NANO 2003)*, vol. 2, pp. 40–43, Aug. 2003.

[41] D. L. Smith and C. Mailhiot, "Theory of semiconductor superlattice electronic structure," *Rev. Modern Phys.*, vol. 62, pp. 173–234, Jan. 1990.

[42] Y.-C. Chang and J. N. Schulman, "Complex band structures of crystalline solids: An eigenvalue method," *Phys. Rev. B*, vol. 25, no. 6, p. 3975, 1982.

[43] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Physics*. New York: Dover Publications, 3rd ed., 2003.

[44] G. D. Mahan, *Many-Particle Physics*. New York: Plenum Press, 2nd ed., 1990.

[45] N. Vukmirovic, Z. Ikonic, D. Indjin, and P. Harrison, "Quantum transport in semiconductor quantum dot superlattices: Electron-phonon resonances and polaron effects," *Phys. Rev. B*, vol. 76, no. 24, p. 245313, 2007.

[46] L. E. Henrickson, "Nonequilibrium photocurrent modeling in resonant tunneling photodetectors," *J. Appl. Phys.*, vol. 91, no. 10, pp. 6273–6281, 2002.

[47] M. Luisier, *Quantum transport beyond the effective mass approximation*. PhD thesis, Eidgenössische Technische Hochschule Zürich, 2007.

[48] K. S. Thygesen and A. Rubio, "Conserving GW scheme for nonequilibrium quantum transport in molecular contacts," *Chem. Phys. Lett.*, vol. 77, no. 11, p. 11533, 2008.

[49] U. Aeberhard, "Quantum transport simulation of hot carrier photocurrent generation in quantum well solar cells," *Semiconductor Sci. Tech.*, vol. 34, no. 9, p. 094002, 2019.

[50] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies," *Phys. Rev. B*, vol. 8, no. 34, p. 5390, 1986.

[51] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body Green's-function approaches," *Rev. Modern Phys.*, vol. 74, p. 601, 2002.

[52] T. C. Kubis, *Quantum transport in semiconductor nanostructures.* PhD thesis, Technische Universität München, München, 2009.

[53] M. Büttiker, "Four-terminal phase-coherent conductance,," *Phys. Rev. Lett.*, vol. 57, p. 1761, 1986.

[54] S. Datta, "Nanoscale device modeling: The Green's function method," *Superlatt. Microstruct.*, vol. 28, p. 253, 2000.

[55] R. Venugopal, M. Paulsson, S. Goasguen, S. Datta, and M. S. Lundstrom, "A simple quantum mechanical treatment of scattering in nanoscale transistors," *J. Appl. Phys.*, vol. 93, p. 5613, 2003.

[56] J. A. Vaitkus and J. H. Cole, "Büttiker probes and the recursive Green's function: An efficient approach to include dissipation in general configurations," *Phys. Rev. B*, vol. 97, p. 085147, 2018.

[57] P. Greck, *Efficient calculation of dissipative quantum transport properties in semiconductor nanostructures.* PhD thesis, Technische Universität München, München, 2012.

[58] P. Greck, S. Birner, B. Huber, and P. Vogl, "Efficient method for the calculation of dissipative quantum transport in quantum cascade lasers," *Opt. Express*, vol. 23, p. 6587, 2015.

[59] B. K. Ridley, *Quantum Processes in Semiconductors.* Oxford: Oxford University Press, 5th ed., 2013.

[60] J. S. Blakemore, "Approximation for Fermi-Dirac integrals, especially the function $\mathcal{F}_{1/2}(\eta)$ used to describe electron density in a semiconductor," *Solid-State Electron.*, vol. 25, no. 11, pp. 1067–1076, 1982.

[61] M. Goano, "Series expansion of the Fermi-Dirac integral $\mathcal{F}_j(x)$ over the entire domain of real j and x," *Solid-State Electron.*, vol. 36, no. 2, pp. 217–221, 1993.

[62] F. Bernardini, "Spontaneous and piezoelectric polarization: Basic theory vs. practical recipes," in *Nitride Semiconductor Devices: Principles and Simulation* (J. Piprek, ed.), ch. 3, pp. 49–68, Weinheim: Wiley-VCH Verlag, 2007.

[63] J. A. Gonzalez Montoya, "Full-band NEGF analysis of nanostructured devices for optoelectronics," Master's thesis, Politecnico di Torino, 2018.

[64] N. I. Bochkareva, V. V. Voronenkov, R. I. Gorbunov, A. S. Zubrilov, Y. S. Lelikov, P. E. Latyshev, Y. T. Rebane, A. I. Tsyuk, and Y. G. Shreter, "Defect-related tunneling mechanism of efficiency droop in III-nitride light-emitting diodes," *Appl. Phys. Lett.*, vol. 96, no. 13, p. 133502, 2010.

[65] U. Aeberhard, "Nonequilibrium Green's function picture of nonradiative recombination of the Shockley-Read-Hall type," *Phys. Rev. B*, vol. 99, no. 12, p. 125302, 2019.

[66] A. Schenk, "An improved approach to the Shockley-Read-Hall recombination in inhomogeneous fields of space-charge regions," *J. Appl. Phys.*, vol. 71, no. 7, p. 3339, 1992.

[67] M. G. Pala, C. Grillet, J. Cao, D. Logoteta, A. Cresti, and D. Esseni, "Impact of inelastic phonon scattering in the OFF state of tunnel-field-effect transistors," *J. Comp. Electron.*, vol. 15, p. 1240, 2016.

[68] D. Esseni, M. Pala, P. Palestri, C. Alper, and T. Rollo, "A review of selected topics in physics based modeling for tunnel field-effect transistors," *Semiconductor Sci. Technol.*, vol. 32, p. 083005, 2017.

[69] K. Jandieri, S. D. Baranovskii, O. Rubel, W. Stolz, F. Gebhard, W. Guter, M. Hermle, and A. W. Bett, "Resonant electron tunneling through defects in GaAs tunnel diodes," *J. Appl. Phys.*, vol. 104, p. 094506, 2008.

[70] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, no. 5, p. 835, 1952.

[71] A. Tibaldi, J. A. G. Montoya, M. Vallone, M. Goano, E. Bellotti, and F. Bertazzi, "Modeling infrared superlattice photodetectors: From nonequilibrium Green's functions to quantum-corrected drift diffusion," *Phys. Rev. Applied*, vol. 16, no. 4, p. 044024, 2021.

[72] G. A. M. Hurkx, D. B. M. Klassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, 1992.

[73] A. Schenk, "A model for the field and temperature dependence of Shockley-Read-Hall lifetimes in silicon," *Solid-State Electron.*, vol. 35, no. 11, pp. 1585–1596, 1992.

[74] P. Sarangapani, Y. Chu, J. Charles, G. Klimeck, and T. Kubis, "Band-tail formation and band-gap narrowing driven by polar optical phonons and charged impurities in atomically resolved III-V semiconductors and nanodevices," *Phys. Rev. Applied*, vol. 12, p. 044045, 2019.

[75] F. Piva, C. De Santi, A. Caria, C. Haller, J. F. Carlin, M. Mosca, G. Meneghesso, E. Zanoni, N. Grandjean, and M. Meneghini, "Defect incorporation in In-containing layers and quantum wells: Experimental analysis via deep level profiling and optical spectroscopy," *J. Phys. D*, vol. 54, no. 2, p. 025108, 2021.

[76] F. Bertazzi, M. Moresco, and E. Bellotti, "Theory of high field carrier transport and impact ionization in wurtzite GaN. Part I: A full band Monte Carlo model," *J. Appl. Phys.*, vol. 106, p. 063718, Sept. 2009.

[77] M. Calciati, M. Goano, F. Bertazzi, M. Vallone, X. Zhou, G. Ghione, M. Meneghini, G. Meneghesso, E. Zanoni, E. Bellotti, G. Verzellesi, D. Zhu, and C. Humphreys, "Correlating electroluminescence characterization and physics-based models of InGaN/GaN LEDs: Pitfalls and open issues," *AIP Adv.*, vol. 4, p. 067118, June 2014.

[78] J. Wu, "When group-III nitrides go infrared: New properties and perspectives," *J. Appl. Phys.*, vol. 106, no. 1, p. 011101, 2009.

[79] A. Tibaldi, J. A. G. Montoya, M. G. C. Alasio, A. Gullino, A. Larsson, P. Debernardi, M. Goano, M. Vallone, G. Ghione, E. Bellotti, and F. Bertazzi, "Analysis of carrier transport in tunnel-junction vertical-cavity surface-emitting lasers by a coupled nonequilibrium Green's function-drift-diffusion approach," *Phys. Rev. Applied*, vol. 14, no. 2, p. 024037, 2020.

[80] T. Schmielau and M. F. Pereira Jr., "Nonequilibrium many body theory for quantum transport in terahertz quantum cascade lasers," *Appl. Phys. Lett.*, vol. 95, p. 231111, 2009.

[81] C. Weisbuch, S. Nakamura, Y. Wu, and J. S. Speck, "Disorder effects in nitride semiconductors: Impact on fundamental and device properties," *Nanophoton*, vol. 10, p. 3, 2021.

[82] C. de Falco, E. Gatti, A. L. Lacaita, and R. Sacco, "Quantum-corrected drift-diffusion models for transport in semiconductor devices," *J. Comp. Phys.*, vol. 204, p. 533, 2005.

[83] M. Filoche, M. Piccardo, Y. Wu, C. Li, C. Weisbuch, and S. Mayboroda, "Localization landscape theory of disorder in semiconductors. I. theory and modeling," *Phys. Rev. B*, vol. 95, p. 144204, 2017.

[84] S.-C. Lee and A. Wacker, "Nonequilibrium Green's function theory for transport and gain properties of quantum cascade structures," *Phys. Rev. B*, vol. 66, no. 24, p. 245314, 2002.

[85] A. Wacker, M. Lindskog, and D. O. Winge, "Nonequilibrium Green's function model for simulation of quantum cascade laser devices under operating conditions," *IEEE J. Quantum Electron.*, vol. 19, p. 1200611, 2013.

[86] G. A. Sai-Halasz, R. Tsu, and L. Esaki, "A new semiconductor superlattice," *Appl. Phys. Lett.*, vol. 30, pp. 651–653, 1977.

[87] A. Rogalski, "HgCdTe infrared detector material: history, status and outlook," *Rep. Prog. Phys.*, vol. 68, pp. 2267–2336, 2005.

[88] D. Z. Ting, A. Soibel, L. Höglund, J. Nguyen, C. J. Hill, A. Khoshakhlagh, and S. D. Gunapala, "Type-II superlattice infrared detectors," *Semiconductors and Semimetals*, vol. 84, pp. 1–57, 2011.

[89] A. Rogalski, P. Martyniuk, and M. Kopytko, "InAs/GaSb type-II superlattice infrared detectors: Future prospect," *Appl. Phys. Rev.*, vol. 4, p. 031304, 2017.

[90] I. Vurgaftman, R. Weih, M. Kamp, J. R. Meyer, C. L. Canedy, C. S. Kim, M. Kim, W. W. Bewley, C. D. Merritt, J. Abell, and S. Höfling, "Interband cascade lasers," *J. Phys. D: Appl. Phys.*, vol. 48, p. 123001, 2015.

[91] F. Bertazzi, A. Tibaldi, M. Goano, J. A. Gonzalez Montoya, and E. Bellotti, "Nonequilibrium Green's function modeling of type-II superlattice detectors and its connection to semiclassical approcaches," *Phys. Rev. Applied*, vol. 14, no. 1, p. 01403, 2020.

[92] L. Zeng, Y. He, M. Povolotskyi, X. Liu, G. Klimeck, and T. Kubis, "Low rank approximation method for efficient Green's function calculation of dissipative quantum transport," *J. Appl. Phys.*, vol. 113, no. 21, p. 213707, 2013.

[93] Z. Tian, R. T. Hinkey, R. Q. Yang, D. Lubyshev, Y. Qiu, J. M. Fastenau, W. K. Liu, and M. B. Johnson, "Interband cascade infrared photodetectors with enhanced electron barriers and p-type superlattice absorbers," *J. Appl. Phys.*, vol. 111, no. 2, p. 024510, 2012.

[94] Ioffe Physico-Technical Institute, St. Petersburg, Russia, "Physical properties of semiconductors." http://www.ioffe.ru/SVA/NSM/Semicond/index.html.

[95] T. Andlauer and P. Vogl, "Full-band envelope-function approach for type-II broken-gap superlattices," *Phys. Rev. B*, vol. 80, p. 035504, 2009.

[96] U. Aeberhard, A. Gonzalo, and J. M. Ulloa, "Photocarrier extraction in GaAsSb/GaAsN type-II QW superlattice solar cells," *Appl. Phys. Lett.*, vol. 112, p. 213904, 2018.

[97] A. Tibaldi, F. Bertazzi, M. Goano, R. Michalzik, and P. Debernardi, "VENUS: a vertical-cavity surface-emitting lase electro-opto-thermal numerical simulator," *IEEE J. Quantum Electron.*, vol. 25, p. 1500212, 2019.

[98] T.-Y. Tsai, K. Michalczewski, P. Martyniuk, C.-H. Wu, and Y.-R. Wu, "Application of localization landscape theory and the $k \cdot p$ model for direct modeling of carrier transport in a type II superlattice InAs/InAsSb photoconductor system," *J. Appl. Phys.*, vol. 127, p. 033104, 2020.

[99] K. Y. Choo, S. V. Muniandy, K. L. Woon, M. T. Gan, and D. S. Ong, "Modeling anomalous charge carrier transport in disordered organic semiconductors using the fractional drift-diffusion equation," *Org. Electron.*, vol. 41, p. 157, 2017.

[100] B. V. Olson, J. F. Klem, E. A. Kadlec, J. K. Kim, and M. D. Goldflam, "Vertical hole transport and carrier localization in InAs/$InAs_{1-x}Sb_x$ type-II superlattice heterojunction bipolar transistor," *Phys. Rev. Applied*, vol. 7, p. 024016, 2017.

[101] L.-W. Wang and A. Zunger, "Pseudopotential-based multiband $k \cdot p$ method for $\sim$ 250 000-atom nanostructure systems," *Phys. Rev. B*, vol. 54, pp. 11417–11435, Oct. 1996.

[102] G. Mugny, F. G. Pereira, D. Rideau, F. Triozon, Y.-M. Niquet, M. Pala, D. Garetto, and C. Delerue, "A study of diffusive transport in 14nm FDSOI MOSFET: NEGF versus QDD," in *46th European Solid-State Device Research Conference, ESSDERC 2016, Lausanne, Switzerland, September 12-15*, 2016.

[103] Y.-M. Niquet, V.-H. Nguyen, F. Triozon, I. Duchemin, O. Nier, and D. Rideau, "Quantum calculations of the carrier mobility: Methodology, Matthiessen's rule, and comparison with semi-classical approaches," *J. Appl. Phys.*, vol. 115, p. 054512, 2014.

[104] J. Glennon, F. Bertazzi, A. Tibaldi, and E. Bellotti, "Extraction of mobility from quantum transport calculations of type-II superlattices," *Phys. Rev. Applied*, vol. 19, p. 044045, 2023.

[105] P. C. Klipstein, Y. Livneh, A. Glozman, S. Grossman, O. Klin, N. Snapi, and E. Weiss, "Modeling InAs/GaSb and InAs/InAsSb superlattice infrared detectors," *J. Electron. Mater.*, vol. 43, p. 2984, 2014.

[106] D. Z. Ting, A. Soibel, and S. D. Gunapala, "Hole effective masses and subband splitting in type-II superlattice infrared detectors," *Appl. Phys. Lett.*, vol. 108, p. 183504, 2016.

[107] S. Rott, P. Binder, N. Linder, and G. H. Döhler, "A combined model for miniband and hopping transport in superlattices," *Physica E*, vol. 2, p. 511, 1998.

[108] N. Yoon, C. J. Reyner, G. Ariyawansa, J. M. Duran, J. E. Scheihing, J. Mabon, and D. Wasserman, "Modified electron beam induced current technique for in (Ga)As/InAsSb superlattice infrared detectors," *J. Appl. Phys.*, vol. 122, p. 074503, 2017.

[109] E. Bellotti, F. Bertazzi, A. Tibaldi, J. Schuster, J. Bajaj, and M. Reed, "Disorder-induced degradation of vertical carrier transport in strain-balanced antimony-based superlattices," *Phys. Rev. Applied*, vol. 16, p. 054028, 2021.

[110] M. Reine, J. Schuster, B. Pinkie, and E. Bellotti, "Numerical simulation and analytical modeling of InAs nBn infrared detectors with p-type barriers," *J. Electron. Mater.*, vol. 42, p. 3015, 2013.

[111] M. Reine, B. Pinkie, J. Schuster, and E. Bellotti, "Numerical simulation and analytical modeling of InAs nBn infrared detectors with n-type barriers," *J. Electron. Mater.*, vol. 43, p. 2915, 2014.

[112] D. Z. Y. Ting, A. Soibel, L. Höglund, and S. D. Gunapala, "Theoretical aspects of minority carrier extraction in unipolar barrier infrared detectors," *J. Electron. Mater.*, vol. 44, p. 3036, 2015.

[113] A. Jang, H. J. Lee, Y. H. Kim, H. Jung, P. Bidenko, and S. Kim, "Effect of barrier layer on InAs/GaSb Type-II superlattice nBn detector," in *International Conference on Electronics, Information, and Communication, ICEIC 2021, Jeju, South Korea, January 31 - February 3*, 2021.

[114] A. Schenk and P. Aguirre, "TCAD models of the ballistic mobility in the source-to-drain tunneling regime," *Solid-State Electron.*, vol. 157, p. 1, 2019.

[115] P. Enders, A. Bärwolff, M. Woerner, and D. Suisky, "$\mathbf{k}\cdot\mathbf{p}$ theory of energy bands, wave functions, and optical selection rules in strained tetrahedral semiconductors," *Phys. Rev. B*, vol. 51, p. 16695, 1995.

[116] R. G. Veprek, S. Steiger, and B. Witzigmann, "Ellipticity and the spurious solution problem of $k \cdot p$ envelope equations," *Phys. Rev. B*, vol. 76, no. 16, p. 165320, 2007.

[117] P.-F. Qiao, S. Mou, and S. L. Chuang, "Electronic band structures and optical properties of type-II superlattice photodetectors with interfacial effect," *Optics Express*, vol. 20, no. 2, pp. 2319–2334, 2012.

[118] G. Liu and S.-L. Chuang, "Modeling of Sb-based type-II quantum cascade lasers," *Physical Review B. Condensed Matter and Materials Physics*, vol. 65, p. 165220, 2002.

[119] E. H. Li, "Material parameters of InGaAsP and InAlGaAs systems for use in quantum well structures at low and room temperatures," *Physica E*, vol. 5, no. 4, pp. 215–273, 2000.

[120] R. G. Veprek, S. Steiger, and B. Witzigmann, "Operator ordering, ellipticity and spurious solutions in $k \cdot p$ calculations of III-nitride nanostructures," *Opt. Quantum Electron.*, vol. 40, pp. 1169–1174, Nov. 2008.

[121] P. Rinke, M. Winkelnkemper, A. Qteish, D. Bimberg, J. Neugebauer, and M. Scheffler, "Consistent set of band parameters for the group-III nitrides AlN, GaN, and InN," *Phys. Rev. B*, vol. 77, no. 7, p. 075202, 2008.

[122] P. Hyldgaard, S. Hershfield, J. H. Davies, and J. W. Wilkins, "Resonant tunneling with an electron-phonon interaction," *Ann. Phys.*, vol. 236, no. 1, pp. 1–42, 1994.

List of publications

Journal papers

1. J. A. Gonzalez Montoya, A. Tibaldi, C. De Santi, M. Meneghini, M. Goano, and F. Bertazzi. “Nonequilibrium Green’s Function Modeling of Trap-Assisted Tunneling in $In_xGa_{1-x}N/GaN$ Light-Emitting Diodes,” *Phys. Rev. Applied*, vol. 16, pp. 044023, 2021.

2. A. Tibaldi, J. A. Gonzalez Montoya, M. Vallone, M. Goano, E. Bellotti, and F. Bertazzi. “Modeling Infrared Superlattice Photodetectors: From Nonequilibrium Green’s Functions to Quantum-Corrected Drift Diffusion,” *Phys. Rev. Applied*, vol. 16, pp. 044024, 2021.

3. A. Tibaldi, J. A. Gonzalez Montoya, M. Alasio, A. Gullino, A. Larsson, P. Debernardi, M. Goano, M. Vallone, G. Ghione, E. Bellotti, and F. Bertazzi. “Analysis of Carrier Transport in Tunnel-Junction Vertical-Cavity Surface-Emitting Lasers by a Coupled Nonequilibrium Green’s Function-Drift-Diffusion Approach,” *Phys. Rev. Applied*, vol. 14, pp. 024037, 2020.

4. F. Bertazzi, A. Tibaldi, M. Goano, J. A. Gonzalez Montoya, and E. Bellotti. “Nonequilibrium Green’s Function Modeling of type-II Superlattice Detectors and its Connection to Semiclassical Approaches,” *Phys. Rev. Applied*, vol. 14, pp. 014083, 2020.

5. A. Tibaldi, J. A. Gonzalez Montoya, F. Bertazzi, M. Goano, M. Daubenschüz, R. Michalzik, and P. Debernardi. “Bridging scales in multiphysics VCSEL modeling,” *Opt. Quantum Electron.*, vol. 51, pp. 231, 2019.

Conference proceedings

1. A. Tibaldi, J. A. Gonzalez Montoya, M. Goano, F. Bertazzi, and E. Bellotti. “Nonequilibrium Green’s function modeling of infrared superlattice photodetectors,” in *2022 US Workshop on the Physics & Chemistry of II-VI Materials*, Oct 2022, Tampa (Florida).

2. J. A. Gonzalez Montoya, A. Tibaldi, C. Casu, N. Roccato, C. De Santi, M. Meneghini, M. Goano, and F. Bertazzi. "Probing carrier transport and recombination processes in dichromatic GaN-based LEDs: a nonequilibrium Green's function study," in *International Workshop on Nitride Semiconductors (IWN 2022)*, Oct 2022, Berlin (Germany).

3. V. Torrelli, J. A. Gonzalez Montoya, A. Tibaldi, P. Debernardi, A. Simaz, M. A. Belkin, M. Goano, F. Bertazzi. "Modeling carrier transport in mid-infrared VCSELs with type-II superlattices and tunnel junctions," in *22th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)*, Sept 2022, Turin (Italy).

4. J. A. Gonzalez Montoya, A. Tibaldi, N. Roccato, C. De Santi, M. Meneghini, M. Goano, and F. Bertazzi. "Trap-Assisted Tunneling in GaN-Based LEDs: A Nonequilibrium Green's Function Study", in *GaN Marathon 2022*, June 2022, Venice (Italy).

5. A. Tibaldi, A. Gullino, J. A. Gonzalez Montoya, M. G. C. Alasio, A. Larsson, P. Debernardi, M. Goano, M. E. Vallone, G. Ghione, E. Bellotti, and F. Bertazzi. "Modeling Tunnel Junctions for VCSELs: A Self-Consistent NEGF-DD Approach," in *20th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)*, Sept 2020, Online.

6. F. Bertazzi, A. Tibaldi, J. A. Gonzalez Montoya, M. Goano, and E. Bellotti. "NEGF Modeling of Carrier Transport in Antimonide-based Type-II Superlattice Absorbers," in *2019 Photonics & Electromagnetics Research Symposium (PIERS)*, June 2019, Rome (Italy).

7. A. Tibaldi, J. A. Gonzalez Montoya, F. Bertazzi, M. Goano, and P. Debernardi. "Bridging scales in multiphysical VCSEL modeling," in *18th International Conference on Numerical Simulation of Optoelectronic Devices (NUSOD)*, Nov 2018, Hong Kong (China).