LUND UNIVERSITY

Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities

Winge, David

2016

Link to publication

Citation for published version (APA):
Winge, D. (2016). *Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities*. Lund University, Faculty of Science, Department of Physics, Division of Mathematical Physics.

Total number of authors:
1

Creative Commons License:
CC BY

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

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

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities

by David Winge

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Andreas Wacker, Dr. Peter Samuelsson
Faculty opponent: Prof. Jacob B. Khurgin

To be presented, with the permission of the Faculty of Science of Lund University, for public criticism in the Rydberg lecture hall (Rydbergsalen) at the Department of Physics on Friday, the 25th of November 2016 at 13:15.

Organization	Document name
LUND UNIVERSITY Department of Physics Box 118 221 00 LUND Sweden	**DOCTORAL THESIS** Date of disputation 2016-11-25 Sponsoring organization
Author(s) David Winge	

Title

Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities

Abstract

This thesis adresses modeling of quantum cascade lasers using non-equilibrium Green's functions. Focus lies on the quantitative modeling of the gain when the systems are exposed to strong laser fields, as under realistic opertional conditions. In the first part, an introduction to the general concepts is provided, together with a discussion on approximate analytic expressions. The second part consists of the six papers listed below.

Paper I analyzes the temperature dependence of the gain and current of a terahertz quantum cascade laser.

Paper II provides details of the implementiation of the non-equilibriun Green's function formalism for simulations of transport in quantum cascade lasers.

Paper III demonstrates a microscopic approach to simulate the second harmonic generation of a mid-IR quantum cascade laser.

Paper IV documents the implementation of a rudimentary form of the GW-approximation for the electron-electron scattering self-energy.

Paper V analyzes a proposed gain mechanism in superlattices, and presents an optimized structure plausible for laser operation.

Paper VI shows simulation results of 16 samples from 5 different laboratories. From this systematic study, trends with repspect to the growth place are identified.

Key words

non-linear response, quantum cascade lasers, non-equilibrium Green's function theory

Classification system and/or index terms (if any)

Supplementary bibliographical information	Language
	English
ISSN and key title	ISBN
	978-91-7753-046-6 (print) 978-91-7753-047-3 (pdf)
Recipient's notes	Number of pages: 162 / Price
	Security classification

Distributor

David Winge, Division of Mathematical Physics, Department of Physics, Sölvegatan 14A, SE-223 62 Lund, Sweden

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____ Date 2016-11-01

DOKUMENTDATABLAD enl SIS 61 41 21

Quantitative Modeling of Gain in Quantum Cascade Lasers under Operational Intensities

by David Winge

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

Cover illustration: Coherence in Superlattice. Adapted from Fig. 4 of Paper V.

Funding information: The thesis work was financially supported by the Swedish Research Council (VR).

© David Winge 2016
Paper I ©2013 by Springer
Paper II ©2013 by IEEE
Paper III ©2014 by Optical Society of America
Paper IV published 2016 under licence (CC-BY) by IOP Publishing Ltd
Paper V published 2016 under licence (CC-BY) by American Institute of Physics
Paper VI ©2016 by American Institute of Physics

Faculty of Science, Department of Physics
ISBN: 978-91-7753-046-6 (print)
ISBN: 978-91-7753-047-3 (pdf)
ISSN: <ISSN number>

Printed in Sweden by Media-Tryck, Lund University, Lund 2016

Acknowledgments

I owe my deepest gratitude to my supervisor Andreas Wacker. Thank you for bringing me to Austria to meet the devouted and inspiring quantum cascade laser community, for introducing me to your colleques and showing me all the interesting work that could be done. Thank you for letting me choose my own way, and for supporting me in everything I tried. I highly value all the things that you have taught me, and all the things that you have helped me teach myself.

Doing my PhD closely together with Martin Frankié has made the studies and the research more fun, much easier, and I think, better. Thank you for sharing thoughts and ideas, for discussing any problem or question, and for inviting me to your and Fridas place for the best soup. Being a part of our bigger research group has been a real pleasure. Thank you Bahareh, Fikkeraddis, Gediminas and Tim, and I wish Kevin the best of luck in Gothenburg! I would like to thank my co-supervisor Peter Samuelsson, for keeping an eye on me these four years. My sincerest thanks to Claudio, for providing valuable discussions on many-body physics.

I have had a long and fruitful collaboration with Mauro F. Pereira at Sheffield Hallam University, and we always manage to have a good time when we meet. My warm thanks to Dragan and Andrew for inviting to me to Leeds University, and for taking very good care of me during my short visit.

I would like to show my gratitude to the seniors at our division who invest a lot of time in preparing suitable courses for us PhD students. This makes a true difference for us.

Thank you Gunnar, for following the tradition of Francesc and Elife in creating the

including and loving atmosphere among the students on our floor, and all of our students, for their wholehearted participation in this endeavor. This place would not be the same without Katarina, whose tireless care of all of us makes this floor like a second home.

I would not have enjoyed my time as a PhD student as much if it were not for my old friends from the Physics department, Johannes, Sanna, Fredrik and Jon. You will always be close to my heart, together with friends like Hamse and André. I also want to thank my great parents, Berit and Harald, for always being only a phonecall away, if only so to start my sour car. Thank you Lena, for being a hero when I need it, and for being it all the other times as well. For all your support, and all your love. Even though Malmö can be rough this time of year, together we keep out the cold.

Populärvetenskaplig sammanfattning

Ljus består av elektromagnetiska vågor som pulserar genom universum med olika frekvenser. Vanligt ljus, som det som kommer från solen och från våra lampor, innehåller ett spektrum av många olika frekvenser. De olika vågornas toppar och dalar svänger inte heller i takt, utan är förskjutna i förhållande till varandra. I laserljus däremot, svänger alla vågor samtidigt upp och ner och med samma frekvens. Vanligt ljus kan liknas vid de kaosartade applåderna efter ett bejublat framträdande, medan laserljuset är det taktfasta klappandet för att få till ett extranummer. Användningsområdena för laserteknik blir hela tiden fler och fler, och idag finns lasern både i dvd-spelaren, i metallindustrin och på ögonklinken.

Våra ögon har förmågan att uppfatta en viss del av ljusets spektrum, från de djuprödaste färgerna till de violetta nyanserna. Andra arter är mer receptiva; näsgropsormen kan uppfatta infraröda signaler och tornfalken kan upptäcka ultravioletta reflektioner från gnagares urin på marken. Idag kan vi producera laserljus i många frekvensområden, men ännu inte i hela det elektro-magnetiska spektrumet. En illustration av detta kan ses i kappan till denna avhandling, i figur 1.1. Ett utvecklingsområde är ett frekvensområde som kallas *Terahertz* (THz), som ligger en bit nedanför det infraröda som näsgropsormen kan uppfatta. En laser i detta frekvensområde kan användas för att leta syre i rymden, scanna väskor och kläder efter vapen och för ultrasnabb spektroskopi. Här sker nu en snabb utveckling tack vare *kvantkaskadlasern*, som är central i denna avhandling. Denna laser består inte utav en gas av atomer eller molekyler, utan av ett solitt system av legereringar. Materialen som används liknar de som används vid tillverkning av LED-lampor och solceller, som genererar eller samlar vanligt ljus. Precis som i LED-lampan,

så är principen att elektrisk energi i form av en ström omvandlas till ljus. Kvantkaskadlasern producerar dock laserljus i THz-området, medan LED-tekniken är konstruerad för vanligt ljus i det synliga området.

I Figur 1.10 i kappan, visas en bild av en klyvfrukt från en lönn, tagen med en THz laser vid Massachusetts Institute of Technology. Andra system har också använts av konservatörer tilsammans med fysiker vid Statens Museum for Kunst i Köpenhamn, för att analysera målningar som är gömda under flera lager av färg. För att kunna göra dessa applikationer mer tillgängliga, och möjliggöra nya, så hoppas man kunna producera en laser i detta frekvensområde som kan fungera vid rumstemperatur. De som finns idag klarar inte detta krav utan måste kylas ned, till exempel med flytande kväve. För att förstå och förbättra dem, krävs en detaljerad förståelse av de mikroskopiska processer som kvantkaskadlasern bygger på. De elektroner som driver kvantkaskadlasern rör sig i strukturer på nanometerskalan, och deras rörelse styrs därför av kvantmekaniska lagar. Istället för att se dem som partiklar måste vi behandla dem som vågor, utan en bestämd plats i rummet. Denna avhandling visar på hur vi med hjälp av kvantmekaniska simuleringar kan beskriva i detalj hur elektronerna i de nanometer-stora systemen beter sig och hur de samverkar med ljusvågorna i strukturerna. Vi undersöker hur övertoner av ljus uppstår i starka laserfält, hur THz-lasrar påverkas av temperetur och hur elektronerna sprids efter kollisioner. Vi utvärderar alternativa idéer i THz-området, och vi jämför slutligen våra beräkningar med experiment från flera världsledande laboratorier. Vår slutsats är att vår modell kvantitativt kan beskriva den kvantmekanska verkligheten i dessa mikroscopiska strukturer.

Popular science summary

Light consists of electro-magnetic waves pulsing through the universe at different frequencies. Ordinary light, like the light originating from the sun and our light bulbs, contain a spectrum of many different frequencies. The peaks and the valleys of the waves are not in phase, but shifted relative to each other. In laser light however, all waves oscillate in a synchronized fashion, up and down, with the same frequency. We can compare ordinary light to the chaotic applause after a great performance, while laser light is more like the steady beat when the crowd is chanting and hoping for an extra set. The areas where laser technology is applied are constantly increasing. Today we find the laser in the DVD-player, in the metal industry and at the eye clinic.

Our eyes can only see parts of the electro-magnetic spectrum, from the deepest red to the violet tones. Other species are more perceptive; the pit viper can sense infrared radiation and the kestrel uses ultraviolet reflections from the urine of small rodents when hunting. Today, laser light can be produced for many frequency bands, but we do not yet cover all of the electro-magnetic spectrum. This is illustrated in the introduction of the thesis, in Figure 1.1. One area of development is a band located a bit below the infra-red that the pit viper perceives, namely the *Terahertz* (THz) band. A laser in this frequency band can be used to look for oxygen in outer space, to scan bags and clothes for weapons and to do ultra-fast spectroscopy. Here, there has been a rapid development thanks to the *quantum cascade laser*, which is central in this thesis. This laser consists not of a gas of atoms or molecules, but instead of solid state alloys. It is constructed of semiconductor materials similar to those used to manufacture LEDs and solar cells, that generates or gather ordinary light. Just as in the LED, electric energy is converted into light. The quantum cascade laser however, emits laser light in the THz band, while the

LED gives ordinary light in the visible.

In Figure 1.10 in the introduction, an image of a dried seed pop of a maple tree is shown, taken with a THz laser at the Massachusetts Institute of Technology. Other systems have been used by conservators in collaboration with physicists at Statens Museum for Kunst in Copenhagen, to investigate artwork hidden under other layers of paint. To make these applications easier, and to make other possible, there is a hope to be able to fabricate a laser in the THz band capable of room temperature operation. Those that exist today does not comply with this demand and needs to be cooled, for example by liquid nitrogen. To understand and improve them, a detailed understanding of the microscopic processes of the quantum cascade laser is needed. The electrons driving the quantum cascade laser transverse structures on the nanometer scale, and their movement is therefore governed by the laws of quantum mechanics. Instead of the particle picture, we have to treat them as waves, with no distinct position in space. This thesis shows how we, with the help of quantum mechanical simulations, can describe in detail how the nanometer scale systems behave and interact with the light waves of the laser structures. We study how higher harmonics of light form in strong laser fields, how THz lasers are affected by temperature and how the collisions of the charged electrons affect the operation. We assess alternative approaches in the THz-band and finally we compare our results with many of the world leading laboratories. Our conclusion is that our model quantitatively can describe the quantum mechanical reality of these semiconductor structures.

Contents

List of publications and author's contributions 1

I Methods and Background 5

1 Introduction 7
- 1.1 Lasers . . . 8
- 1.2 Superlattices . . . 10
- 1.3 Quantum cascade lasers . . . 14
- 1.4 Challenges to theory . . . 18

2 Optical response of a medium 21
- 2.1 Propagation of light . . . 21
- 2.2 Susceptibility through time-dependent perturbation theory . . . 23
- 2.3 Our approach to the susceptibility . . . 29

3 Non-equilibrium Green's functions 31
- 3.1 Motivation . . . 31
- 3.2 Theoretical background . . . 33
- 3.3 Implementation . . . 37
- 3.4 Frequency expansion in the Green's function . . . 41
- 3.5 Electron-electron scattering . . . 42
- 3.6 Gain saturation . . . 49
- 3.7 Summary . . . 50

4 Quantitative results from non-linear response simulations 51
- 4.1 Temperature dependent gain saturation . . . 51
- 4.2 Quantitative results under operation . . . 52
- 4.3 Simulating second harmonic generation . . . 57

5 Revealing physics in quantum cascade lasers 61
- 5.1 Second harmonic generation from perturbation theory . . . 61

5.2 Gain characterization 63
5.3 Parametrization of the QCL response 66

6 Summary & Outlook 73

Acronyms 75

References 77

II The papers 89

List of publications and author's contributions

This thesis is based on the following publications, referred to by their Roman numerals:

I **Temperature dependent nonlinear response of quantum cascade structures**

D. O. Winge and A. Wacker
Optical and Quantum Electronics **46**(4), pp. 533–539 (2013)
©2013 by Springer

In this publication the current and gain of a terahertz quantum cascade laser was studied under different temperatures. The results showed qualitative agreement when compared to experimental data. The simulated maximum operating temperature was found to be in good agreement with other studies on similar designs.

For this paper I produced the data and figures and wrote a first draft.

II **Nonequilibrium Green's Function Model for Simulation of Quantum Cascade Laser Devices Under Operating Conditions**

A. Wacker, M. Lindskog, and **D. O. Winge**
IEEE Journal of Selected Topics in Quantum Electronics **19**, p. 1200611 (2013)
©2013 by IEEE

This paper provided an detailed description of our implementation of the non-equilibrium Green's function formalism for quantum cascade laser simulations, including the most recent changes.

Here I produced the data for Figs. 5-7 and wrote a draft for the sections on *High intensity* and *Integrated absorption* (Sec IV. B and C) as well as a draft for the Appendix.

III **Microscopic approach to second harmonic generation in quantum cascade lasers**

D. O. Winge, M. Lindskog, and A. Wacker
Optics Express **22**, pp. 18389-18400 (2014)
©2014 by Optical Society of America

In this work a mid-IR quantum cascade laser optimized for second harmonic generation was studied. From our computational model, the intensity both at the fundamental and the double frequency could be calculated and compared to the experiment with good agreement.

For this paper I came up with the idea, produced the data and figures and wrote a first draft.

IV **Simple electron-electron scattering in non-equilibrium Green's function simulations**

D. O. Winge, M. Franckié, C. Verdozzi, A. Wacker, and M. F. Pereira
Journal of Physics: Conference Series **696**, p. 012013 (2016)
Published 2016 under license (CC-BY) by IOP Publishing Ltd

This paper documented the implementation of a rudimentary form of the GW-approximation for the electron-electron scattering self-energy.

For this project I did the derivations together with M.F. Pereira, wrote the code calculating the new self-energies, produced the data and figures and wrote a first draft.

v **Superlattice gain in positive differential conductivity region**

D. O. Winge, M. Franckié, and A. Wacker
AIP Advances **6**, 045025 (2016)
Published 2016 under license (CC-BY) by American Institute of Physics

In this paper a proposal for superlattice gain in a region of positive differential conductivity was assessed. An in depth analysis of the gain showed that the dipole matrix elements were strongly dependent on the scattering. Plausible laser operation at low temperatures was suggested for an improved structure.

For this paper I produced parts of the data and all the figures, performed the analysis of the matrix elements discussed in the second half of the paper and also wrote parts of a first draft.

vi **Simulating terahertz quantum cascade lasers: Trends from samples from different labs**

D. O. Winge, M. Franckié, and A. Wacker
Journal of Applied Physics **120**, p. 114302 (2016)
©2016 by American Institute of Physics

A large study of 16 samples was presented. The samples, being from five different laboratories, allowed us to reveal trends in the measured current with respect to the origin. At the same time, the model was shown to provide consistent results for samples from the same growth place.

For this project I modified the code to improve the calculation of the self-energies. I contributed with half of the simulation data that make up the bulk of the results shown in Tab. 1. I produced the data for Figs. 1-2 and 5-6 and made all figures. I wrote a first draft for the manuscript.

All papers are reproduced with permission of their respective publishers.

Other publications, not included in this thesis:

Nonlinear response of quantum cascade structures

D. O. Winge, M. Lindskog, and A. Wacker
Applied Physics Letters **101**, p. 211113 (2012)

This work has a large overlap with the content of the diploma thesis for my M.Sc. degree, and is therefore not included here.

Injection Schemes in THz Quantum Cascade Lasers Under Operation

M. Lindskog, **D. O. Winge**, and A. Wacker
Proc. SPIE, 8846:884603 (2013)

Impact of interface roughness distributions on the operation of quantum cascade lasers

M. Franckié, **D. O. Winge**, J. M. Wolf, V. Liverini, E. Dupont, V. Trinité, J. Faist, and A. Wacker
Optics Express **23**, p. 5201-5212 (2015)

Influence of Interface Roughness in Quantum Cascade Lasers

K. Krivas, **D. O. Winge**, M. Franckié, and A. Wacker
Journal of Applied Physics **118**, p. 114501 (2015)

Part I

Methods and Background

1 Introduction

Today, the laser is an important light source in most areas of science and research. Far from the visible part of the electromagnetic spectrum, where laser pointers as shown in Fig. 1.3 are now commonplace gadgets, lasers are also available from the ultraviolet region with short wavelengths, to the infra-red part with long wavelengths. In Fig. 1.1, it is shown how well commercially available lasers of different types together cover a large range of the electro-magnetic spectrum. Short wavelengths are practical in imaging and reading if a high resolution is desirable, as in Blu-ray and DVD players, and allow for a high density of data on the discs. On the contrary, long wavelengths are lower in energy, and these rays can penetrate deeper into materials, like tissue or cloth. This makes them usable for medical applications and spectroscopy, as many important trace gases like carbon dioxide and methane, have strong absorption lines in the low wavelength region. This thesis concerns the rightmost part of this figure, where the usual semiconductor lasers, molecular gas lasers and chemical lasers find it hard to extend their range to lower frequencies [1]. Here the coverage of the *quantum cascade laser* is indicated by a navy blue color. In this thesis, the physics of this intriguing laser device is studied, with a focus on microscopic simulations of light generating mechanisms, all on the nanometer scale.

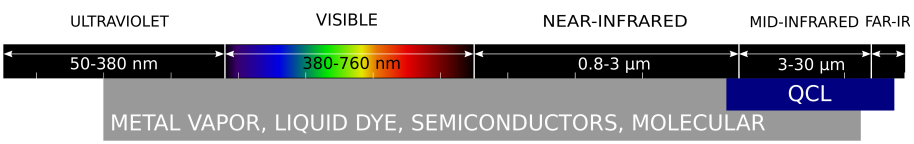

Figure 1.1: Commercial laser coverage of the electro-magnetic spectrum indicated by the gray field. The region covered by the quantum cascade laser (QCL) is marked by a navy blue region. A few examples of the important laser types are given, for a full coverage, please consult Ref. [1].

To introduce this thesis we will start by discussing the solid state systems that we are interested in. This introduction will serve to familiarize the reader with the way we think about these structures and how we present them, and we will also try to motivate the importance of this research field, both in terms of basic research in physics and in technological applications. Then, in Chap. 2, we will discuss some of the physical observables that we want to calculate and understand. After that, in Sec. 3, we present our method and the theory that underlies it. In Sec. 4 the quantitative achievements based on our model are presented and in Sec. 5 we interpret our results to identify the underlying physical phenomena, before we focus on a few conclusions and look forwards in Sec. 6. The philosophy and the pedagogic aspect of this way of presenting the thesis, is that the beginning can read by all, and depending on how interested the reader is in the actual details of the method and the specific results of our different studies, one can choose to read more or less of the more advanced parts that will be located in the end of the thesis.

1.1 Lasers

The fantastic and very useful devices that are called lasers carries a short hand name for something slightly more complicated. The acronym LASER stands for *light amplified by stimulated emission of radiation*.

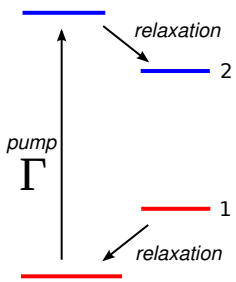

Figure 1.2: Four-well laser scheme with a pumping rate Γ.

Stimulated emission is possible in a system where electrons are excited, from a low energy state to a higher energy state. These electrons, having a surplus of energy, are able to transfer their excitation energy to a light field. Let us consider the four-level system depicted in Fig. 1.2, where four possible energy states for electrons are shown, ordered in energy. At thermal equilibrium, most electrons will be in the lowest possible energy state, the ground state. If a light field, consisting of photons matching the energy gap of the lowest and highest energy state, a process of *stimulated absorption* would take place, where photons are absorbed by the electrons which then gets excited to the highest energy state. From there, searching for ways to loose energy and to get back to the ground state, they eventually fall down

to state 2 as indicated in Fig. 1.2, for example by scattering processes. If the probability of scatter from state 2 to state 1 is low, the electrons will stay excited for a considerable time. Apart from scattering, they can emit photons and thereby lose enough energy to deexcite to state 1. This is a slow process called *spontaneous emission*. However, if there are already photons in the system with energy matching the transition, the quicker process of *stimulated emission* is possible, where the electrons deexcite and produce a photon identical to the photon that "encouraged" the emission. Here, it is important to emphasize that the generated photon will have the exact same frequency and phase as the first photon.

From this discussion and with the help of Fig. 1.2, we summarize the principle to generate laser light [2]. Electrons are pumped into a high energy state by a rate Γ, from where they quickly relax into the upper laser state (state 2 in Fig. 1.2). This creates an *inverted* population, i.e. more electrons in state 2 than in state 1, as opposite to the situation in equilibrium. A photon in the system matching the transition energy are now more likely the get an identical copy than being absorbed. This is called *optical gain* and it is the opposite to absorption. To protect the inverted situation, it is important to have a quick depopulation of the lower laser state (state 1 in Fig. 1.2).

As photons travel with the speed of light, it is important to trap them to enable the stimulated emission. This is done by placing the system in between two mirrors, allowing power to build up. In a real system, there will also be losses, for example by imperfect mirrors where photons can leak out. Lasing first starts when the optical gain can overcome the losses. At this point, there will be an avalanche-like effect when the photons in the cavity rapidly multiply. To summarize the above discussion, an efficient laser employs fast injection of the upper laser state, quick depopulation of the lower laser state and confines the light field between two mirrors.

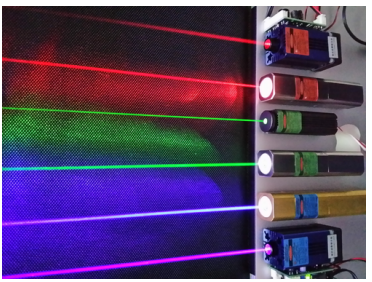

Figure 1.3: Solid state lasers at different optical wavelengths. Photo by Pang Ka kit [3] (CC BY 2.5).

The work presented here aims to quantitatively calculate the optical gain of the devices. These are solid state devices that are electrically pumped, with electrons being injected to the upper laser state by a current [4]. As the physical processes,

giving the optical gain in quantum cascade lasers (QCLs), are electronic transitions between states very close in energy (THz QCLs operate with energies 0.07% of the hydrogen binding energy), accurate modeling needs to account for all processes relevant at this energy scale, which are quite a few. In contrast to the electrically pumped laser diodes (a few examples shown in Fig. 1.3), working roughly like a reversed solar cell, the QCLs employs a cascading principle which recycles the electrons, and injects them from the ground state of one laser period, directly to the upper laser state of the next period. Below, we will see in detail how the principles for generating laser light can be employed in solid state systems.

1.2 Superlattices

The simplest system considered in quantum mechanics is the quantum well, where a particle is confined between two walls, which results in discrete energy levels. The objective of QCL design is to produce an inverted population of electrons in such a system, by the scheme presented in Fig. 1.2. Quantum wells are commonly realized in semiconductor materials, with one material for the barrier and one material for the well. At low temperature, semiconductors have a completely filled valence band and an almost empty conduction band. This makes the poor conductors, as inferred by their name. To improve the conductivity, dopant atoms with extra electrons are inserted to produce n-doped semiconductors, with more electrons in the conduction band. This provides a convenient way to tune the conductivity. A superlattice (SL) consists of a periodic sequence of quantum wells, depicted in Fig. 1.4. The word can be understood as a lattice of quantum wells, *super*-imposed

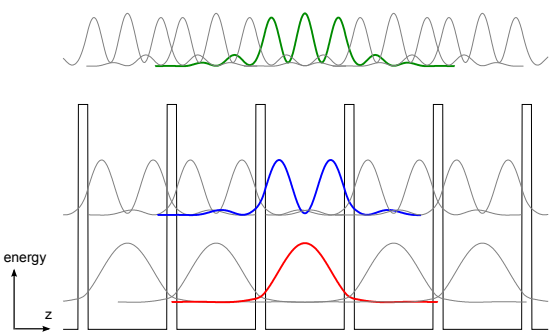

Figure 1.4: A superlattice is normally presented by showing an energy diagram. Here, the jagged conduction band edge is shown together with the discrete states where the electrons are allowed to reside.

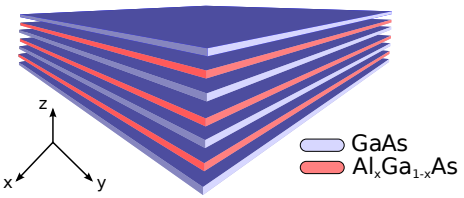

Figure 1.5: Illustration of how semiconductor material is stacked on top of each other during epitaxial growth of heterostructure wafers. Each layer has a different bandgap which can, for example, manifest itself as quantum wells in the valance and the conduction band.

on a lattice of atoms of the material of which the quantum wells are One well can usually fit 100 individual atoms, however, in this thesis we will not study the physics on the atomic level, rather the larger systems of coupled quantum wells. To do this we use *envelope functions* as effective states [5, 6], like the ones shown in Fig. 1.4, that span many atomic sites. The lattice of atoms themselves will be treated as an effective background and enter through several material parameters. A few examples of such parameters are the band gap, the background refractive index, and the effective mass of the electrons.

The quantum wells of Fig. 1.4 are produced via conduction band engineering of semiconductor materials, and a thorough introduction to the physics can be found in Ref. [7]. For a given conduction band offset (CBO) of a combination of two materials, the width of the barriers and wells can be varied. Thicker barriers will give less overlap of the discrete states and less current when we apply a bias over the structure. If the wells are made wider, more discrete states will be allowed to fit into the well. In the SL shown in Fig. 1.4 two states are confined whereas the third is in the *continuum*, that is, above both the well and barrier conduction band edge. This corresponds to the highest state shown here. The states are plotted with their respective probability density at an arbitrary scale, using their discrete level as a baseline.

In the laboratory, the SLs are grown by depositing layers of different material on top of each other. A schematic layer sequence is depicted in Fig. 1.5. The growth direction is along z, where we have quick changes in material composition giving rise to the quantum wells shown in Fig. 1.8. From this schematic view with only a few layers, one can see that the carriers will be confined with respect to the growth direction z, but almost free to move in the other two directions x and y. This is why the band diagrams, as Fig. 1.4, usually include only one direction, as the system is symmetric in the other two directions. When we model the structure, we will

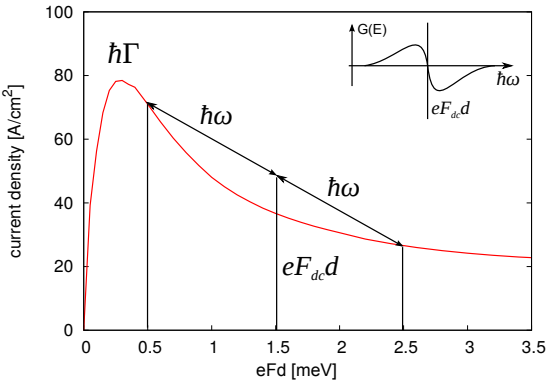

Figure 1.6: Current-voltage (IV) characteristics for the superlattice in Fig. 1.4 for small biases. The first peak occurs when the applied bias per period eFd is equal to the scattering rate $\hbar\Gamma$ in energy units. The inset shows dispersive gain around the Bloch frequency $eF_{\mathrm{dc}}d/\hbar$.

assume that the system is infinite in the x and y direction and take advantage of this rotational symmetry. The layered structure is grown using either metal-organic chemical vapor deposition (MOCVD) or molecular beam epitaxy (MBE), which are techniques able to deposit semiconductor alloys on the substrate, one atom layer at a time.

As this thesis is focused on gain in heterostructures, we also provide a short historical background on optical gain in SLs. Under applied bias, the electrons in the conduction band of the SL will oscillate with the Bloch frequency $eFd/\hbar$, where e is the elementary charge, F is the electric field and d the period length. In a usual material like a metal, the electrons will rarely complete an entire cycle. Instead they get scattered with the a rate $\Gamma/\hbar$ and the oscillation starts over. In the artificial lattice of the SL, the scattering rate Γ can be engineered to be much lower than the Bloch frequency, and it is possible for an electron to complete many cycles before it scatters. This periodic motion can be tapped of energy [8], and used to enhance an optical field [9], a process known as *Bloch gain*.

To illustrate this, we plot the simulated current as a function of bias $I(eFd)$, for our SL of Fig. 1.4 in a new Fig. 1.6. In this zoom, into the low bias region, we clearly observe the so called *miniband peak* predicted by Esaki and Tsu [10]. This occurs when the applied bias is equal to the scattering rate, $eFd = \hbar\Gamma$. This is the field when Bloch oscillations start to affect the system strongly. After this peak, the current goes down when we apply a higher bias. This is called a region of negative differential resistance (NDR) or negative differential conductance (NDC), and the

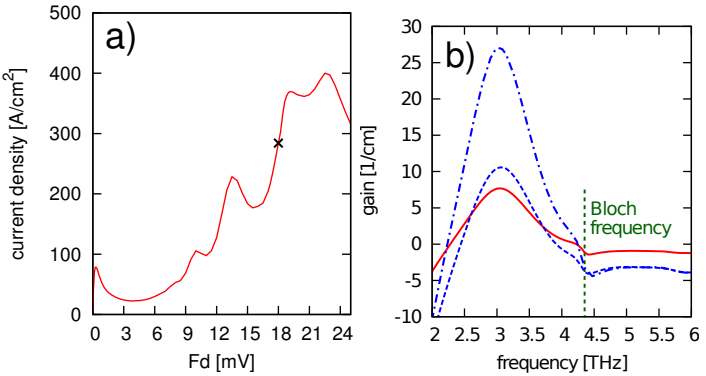

Figure 1.7: Adapted from Paper V. Simulation results for the SL shown in Fig. 1.4 where we plot the IV characteristics in (a) and the gain spectra at 77 K (red) in (b) for the nominal structure at 77 K lattice temperature. In addition, we show the gain of an optimized structure (blue) also at 77 K (dashed) and 40 K (dotted-dashed). The gain is taken at 18 mV per period as indicated by a cross in (a).

Bloch gain is inherently linked to these regions. In order to enhance an electromagnetic field, the slope from $I_{dc}(eF_{dc}d-\hbar\omega)$ to $I_{dc}(eF_{dc}d+\hbar\omega)$ has to be negative [11]. As seen in the figure, this is only possible if the frequency is below the Bloch frequency. The practical problem with the Bloch gain is that if the SL is biased in the NDR, the field distribution will no longer stay homogeneous, but collapse into different electric field domains [11], where each domain is in a region of *positive* differential resistance. So far, no SL devices using the Bloch gain for a lasing has been reported.

In Paper V, we analyze an alternative idea to produce gain in a positive differential region of conductivity of superlattices, in contrast to the Bloch gain. The proposal of Refs.[12, 13] was based on the SL shown in Fig. 1.4. It was modeled by us to quantify a possible gain mechanism and the results are shown in Fig. 1.7(a). In the plot showing current, the miniband peak is again visible at low bias. As the bias of the structure increases, current peaks for each tunneling resonance can be seen. At 18 mV per period, the operating point is reached. Gain was simulated here, and a high signal can be observed around 3 terahertz (THz) in Fig. 1.7(b). In this context, it should be added that ordinary metal-metal (MM) waveguides commonly used for THz QCLs, have losses around 15/cm at these frequencies. There is actually also a small signature around the Bloch frequency that has the proper shape, as shown in the inset of Fig. 1.6, but it is much smaller than the direct inversion gain at 3 THz.

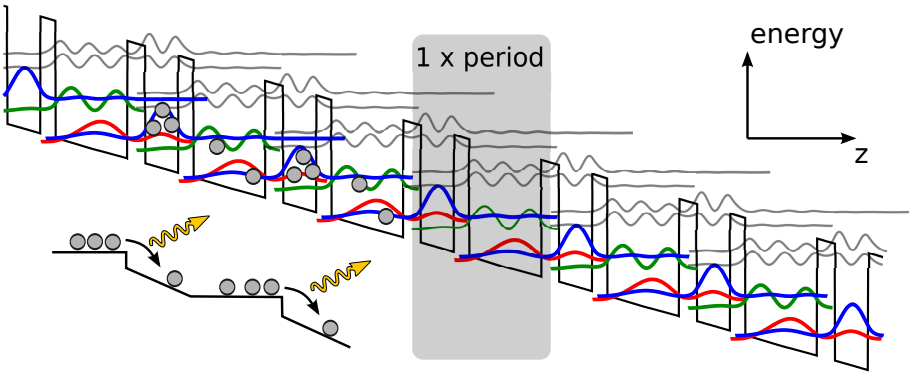

Figure 1.8: An example of a THz QCL from Ref. [15] where each period, as indicated in the figure, consists of two wells and two barriers each. The illustration shows the conduction band edge in black, where the bandgap is engineered to have coupled quantum wells with quantized states. Here, the probability density of these states are plotted with arbitrary scale at their respective energies.

1.3 Quantum cascade lasers

The QCL was first realized in 1994 in the group of Federico Capasso, by Jerome Faist and coworkers at Bell Labs [14]. It is a generalization of the superlattice, where each period consists not of one quantum well only, but instead a sequence of them, again repeated many times. This laser is electrically pumped with carriers injected into the conduction band of a semiconductor heterostructure. Via engineering of the heterostructure, and an applied bias over the structure, these carriers can be pushed to states high in energy; creating a local inversion with respect to lower lying states. This inversion forms the basis for the stimulated emission in the laser cavity.

In Fig. 1.8, a minimal working example of a THz QCL is shown. The structure was demonstrated in Ref. [15]. What is depicted is the conduction band edge (black), where quantum wells are formed using different semiconductor alloys for barriers and wells. In this case, the material system is $Al_xGa_{1-x}As$, where x denotes the aluminum (Al) content versus the gallium arsenide (GaAs). The wells consists of pure GaAs while the barriers have 15% Al. Inversion is produced when the injector state (red) aligns with the upper laser state (ULS) (blue), which is then populated through resonant tunneling. The lower laser state (LLS) (green) is positioned in energy exactly one longitudinal optical (LO) phonon energy ($E_{LO} = 37$ meV in GaAs) above the injector state of the next period. When energetically possible,

the LO phonon scattering is the fastest scattering process in the semiconductors material systems addressed here, and this will ensure a rapid depopulation of the LLS. The energy dependence is important; as the laser transition is below the Reststrahlen band ($\hbar\omega < E_{\mathrm{LO}}$), the inversion is not destroyed by the emission of LO phonons unless the electrons in the ULS are highly excited. In the ideal case, the electrons tunnel into the ULS, and due to the absence of any strong scattering mechanisms, the inversion is sustained long enough for stimulated emission to occur. Here they undergo a transition to the LLS, are quickly deexcited by LO phonon emission to the injector state of the next period, from where they can repeat the process. The bias is doing the pumping in this structure: it supplies the energy needed to lift the electrons that have cascaded down through the total number of periods, up to the first period, to make them start all over again.

The modes of the electromagnetic field have to be supported by the laser cavity. This might sound trivial, but as it is very hard to grow the periodic structure of the QCLs, they are often not thicker than 10 μm. For THz radiation, with wavelengths around 100 μm, this requires waveguides with sub-wavelength confinement[16, 17].

Compared to earlier solid state lasers, which rely on transitions across the bandgap of semiconductor materials of the order of 1 eV and thus have optical transitions around the visible range, the QCL uses transitions within *subbands* in the conduction band, and thereby exhibits much lower transition energies. In addition, while the band gap is fixed, the dimensions of the quantum wells are engineering parameters. This makes it possible to freely tune the transition frequency, i.e. the wavelength of the laser, by only modifying the bandgap profile. Although this might sound simple in theory, there are several issues that need to be solved. Instead of emitting a photon, an electron can relax in a number of different ways. Those transitions, that we will denote *non-radiative*, are possible through scattering from either phonons or imperfections in the structure. We will treat these effects more carefully in Sec. 3 where we discuss the modeling in detail. As the scattering processes are very quick, one has to design these structures so that the carriers are trapped in an upper laser state for a sufficiently long time. Only then can we have inversion and stimulated emission. Equally important, the lower laser state needs to be emptied if lasing is to be maintained. Here, the fast scattering processes can instead be exploited to design an efficient extraction process.

The total number of layers in a QCL is usually over 1500, making up over 200

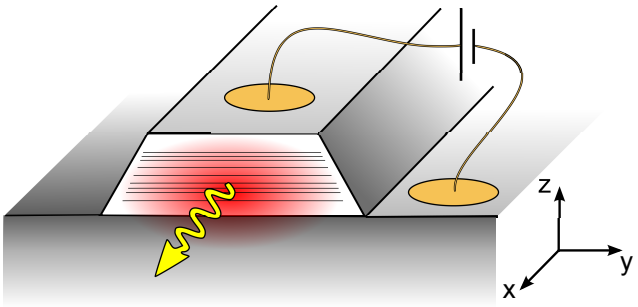

Figure 1.9: Illustration of the actual laser where the wafer has been cleaved and processed with contact layers and connected to a circuit. Under the red area, indicating the waveguide mode, the individual layers of the heterostructure can be seen. Light is emitted preferably along the normal vector of the facet, depending on waveguide design.

periods for THz devices. In total, this gives a height of 10 μm, and the wafer is then cleaved and processed with waveguides and contacts. These contacts are connected to the bottom and the top of the layer sequence as depicted in Fig. 1.9, and when the structure is biased, the structure will experience a potential drop as shown in Fig. 1.8 where the structure was biased with a field of 12 kV/cm. The waveguide facets are usually left untreated; they are nothing more than the cleaved surface of the layered structure. Still they serve as the mirrors of the laser cavity, since the difference in refractive index is high (3.6 for GaAs compared to 1 for air or vacuum). Light will thus propagate back and fourth in the cavity at the round trip time.

There are essentially two different kinds of QCLs: mid-infrared (mid-IR) devices, operating at energies above the Reststrahlenband with wavelengths of 3-25 μm, and THz devices operating at energies below it, with wavelengths from 50-360 μm. The Reststrahlen band is enclosed by the transverse and LO phonon resonances, and here the dielectric function is negative and thus no light can propagate. Below the Reststrahlen band, the transition energies are of the order of 10 meV. These energies correspond to frequencies in the THz range, which makes the QCL a possible candidate to fill what is known as the *THz gap*. This gap constitutes the upper limit of electronic sources, such as backward wave oscillators, Gunn and resonant tunneling diodes oscillators, and Schottky diode multipliers [19, 20]. These microwave and millimeter devices have been developed to cover the GHz range. From above, the gap is at the lower limit of the photonic devices, such as the QCL, and the devices capable of generating THz radiation at room temperature are either optically pumped gas lasers or free electron lasers, or methods that rely

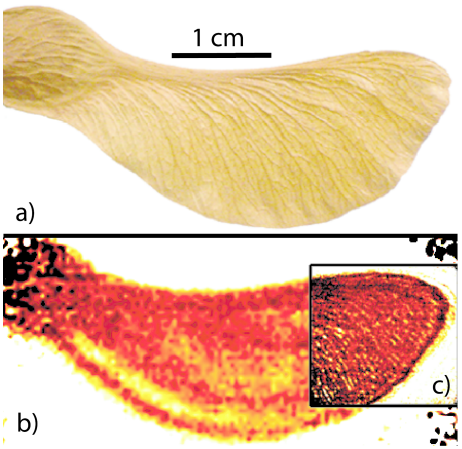

Figure 1.10: A dried seed pod imaged with optical frequencies in a), and with the THz imaging system of Ref. [18] in b) and c) for two different configurations. In both cases, the THz radiation propagated $\sim$ 25 m through air before the hitting the object. The images of b) and c) are both taken with 1 s integration time. Picture courtesy of Alan Lee, reproduced with permissions from [18], Copyright AIP (2006).

on down-conversion of higher frequencies [19]. A successful example of down-conversion is THz sources based on difference frequency generation of two mid-IR lasers [21]. This is currently the most promising direction for using QCLs for the generation of coherent THz radiation at room temperatures [22] and it has been shown to lase in continuous wave mode, although with small output power [23].

The first mid-IR QCL was realized in the InGaAs/AlGaAs material system. It was later also realized in the AlGaAs material system [24, 25], and the mid-IR QCLs are now reliable sources at room temperature [26]. Using distributed feedback gratings, the fabrication of single-mode tunable sources with narrow linewidth is possible. These are ideal for spectroscopic applications, as is discussed in Ref. [27]. The devices are continuously finding new applications, as discussed in Ref. [28]. For example, many important green house gases have spectral fingerprints in this frequency range, which makes the QCL practical for environmental monitoring.

The THz QCL was first demonstrated in 2002 by Kohler *et al.* and coworkers [16] and today it has been shown to lase up to 200 K [29]. At low temperatures, they are high power sources with narrow linewidths, as discussed in Ref. [30], and ideal for use as local oscillators in spectroscopic applications. One very interesting project that is currently using a THz QCL to probe our galaxy, is the SOFIA project. This is a collaboration between the German Aerospace center and NASA, and they use

a 4.7 THz QCL as a local oscillator for heterodyne spectroscopy [31]. The laser is designed to have its frequency around a fine structure line of neutral oxygen, an element of great interest to astronomers, and it is mounted inside a Boeing 747 capable of taking it to a height of 41 000 feet to screen off most of the oxygen of the earth. Another interesting example for security and medical applications, is the possibility to use THz QCLs for imaging [32, 33, 34]. In Fig. 1.10, a picture from Ref. [18], taken with a THz QCL, is shown. By using the transmission window at 4.9 THz, Lee *et. al* could transmit the radiation over 25 m before imaging and collecting the signal. Among the more recent applications, the groups of Fukunaga and Jepsen demonstrated how THz time domain spectroscopy can be used for art conservation [35]. Using this technique together with x-ray radiography, they could extract valuable information about a hidden portrait under an easel painting of the important Danish neoclassical painter Nicolai Abildgaard.

The requirement to cryogenically cool the laser sources hinders the full range of applications. For these, as well as other applications, it is a major goal for the research community to reach room temperature THz QCLs. In order to do this, a thorough understanding of the physics of these intersubband devices is needed.

1.4 Challenges to theory

So why is there still theoretical research on these systems, although they were introduced over 20 years ago? Although great progress has been made in the development of modeling techniques, the review of Ref. [36] provides a thorough summary, there are still many things left to study, especially in the THz region, where room temperature has not yet been demonstrated, and perhaps never will without down-conversion. One important question is how strong the electron-electron scattering is in these structures. Another question, that the modeling community still struggles to answer, is what the main temperature degrading mechanism is; increased scattering, thermal back-filling or even leakage to the continuum [37]? In addition to this, frequency combs have been realized in THz QCLs [38, 39] and the conditions for their formation is still a developing field [40]. These processes include highly non-linear optical effects that we will discuss further in Sec. 2.

The physics of these devices is complicated for a number of reasons. In the band diagram of Fig. 1.8, we can identify a number of scattering mechanisms. Each period is n-doped with another type of atoms to generate carriers in the conduction

band. This introduces impurities which can scatter and dephase the electrons. Furthermore, the interfaces between wells and barriers are not perfect and introduces additional scattering, and so does the mixed alloy in the barriers. As mentioned above, there is also scattering by phonons, which is an inelastic process that can dissipate heat from the electrons. As there is a bias over the structure, the system is out of equilibrium. This implies limited lifetimes of each state and thus a broadening in energy. In addition to scattering mechanisms, the transport of electrons through such a system is inherently related to coherent superpositions of energy eigenstates [41], which needs to taken into account in modeling.

In addition to the requirements above, intersubband structures also exhibit a strong non-linear response [42], which is very interesting to model as the QCLs are high power devices. This requires a strong ac field in the simulations.

To summarize the problems, one can look at the energy scales of the different phenomena, for example in a THz QCL. The width of the states is of the order of 10 meV, the same as the typical transition energy for lasing. Under operation, a typical ac field strength is also of the order of 10 meV. From the theoretical point of view, this tells us that we will have to treat all these physical phenomena to the same order, if we are not to make any systematic errors in our modeling. In Sec. 3, we will introduce and motivate our method of choice for modeling semiconductor heterostructures.

2 Optical response of a medium

A material subjected to a light field can respond in several ways. Depending on the frequency of the light, it may reflect the waves, like metal does at optical frequencies, or let them pass trough almost unscattered, like window glass or air. In this chapter, we discuss first how light propagates through a medium. Then we analyze the optical response of systems consisting quantized states like those in a SL or a QCL, in terms of quantum mechanical operators. We conclude by relating the perturbative expansion in this chapter to how the optical response is treated in the implementation of the non-equilibrium Green's function (NEGF) computational model.

2.1 Propagation of light

In this section we will discuss how light propagates through a medium. If subjected to an electric field, the electrons in a material will respond and react to it by reconfigurating their orbitals. This is called the *polarization* $\mathbf{P}$ of a medium and together with the applied electric field $\mathbf{E}$, this gives the displacement field $\mathbf{D}(\omega) = \epsilon(\omega)\mathbf{E}(\omega)$, where $\epsilon(\omega)$ is the dielectric function tensor as a function of frequency ω. In this chapter, bold letters will denote vectors where we indicate unit vectors with a hat. Consider the electric field polarized in the $\hat{\mathbf{z}}$ direction and traveling along $\hat{\mathbf{x}}$,

$$\mathbf{E}(\mathbf{r}, t) = \hat{\mathbf{z}} F_{\text{ac}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \cos(\omega_0 t) \tag{2.1}$$

where ω_0 is the angular frequency of the field and $\mathbf{k}$ is the wavevector, and the F_{ac} gives the strength of the oscillating field (hence ac for alternating current). The

wavevector can be split in a real and an imaginary part, for the x-component for example as $\tilde{k}_x = k_x + \mathrm{i}\alpha_x/2$, where the tilde indicates a complex number and the imaginary part is the absorption α divided by two. To understand the meaning of the factor one half, consider the absolute square of Eq. (2.1), which is proportional to the intensity. Only the imaginary part of the exponential survives, and reads $\mathrm{e}^{-\alpha x}$ with our definition of the absorption, which gives a natural measure of the *intensity absorption* per unit length.

In a material, the wave vector and frequency are related by the *dispersion* as

$$\tilde{k}_x = \frac{\omega\sqrt{\epsilon_{xx}(\omega)}}{c} = \frac{\omega\tilde{n}_{xx}(\omega)}{c} \tag{2.2}$$

where $\tilde{n}(\omega)$ is the complex refractive index, linked to the relevant component in the three dimensional dielectric tensor. In the following we will assume an isotropic medium, which means that the dielectric tensor is diagonal ($\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz}$), and drop the indexes when not necessary. If the dielectric function is negative, we see that the refractive index $\tilde{n}(\omega)$ becomes purely imaginary and light cannot propagate, as the exponential in Eq. (2.1) dies off exponentially. This is what happens in the Reststrahlen band of semiconductors.

The polarization is of a medium from small perturbations is given by the *relative susceptibility* $\chi^{\mathrm{rel}}(\tau)$, a response function describing the memory of the system as

$$\mathbf{P}(t) = \epsilon_r\epsilon_0 \int_{-\infty}^{t} \mathrm{d}t' \chi^{\mathrm{rel}}(t-t')\mathbf{E}(t'), \tag{2.3}$$

where we define the real relative permittivity ϵ_r following Ref. [43], taking into account contributions from the background. This could be, for example, a crystal lattice with no dopant atoms present. As discussed in Chap. 1, we study only the envelope functions in this thesis, ranging over many atomic orbitals. The underlying material system enters through parameters like ϵ_r. We similarly have the dielectric field tensor $\epsilon(\omega) = \epsilon_0\epsilon_r\epsilon^{\mathrm{rel}}(\omega)$ with $\epsilon^{\mathrm{rel}}(\omega) = 1 + \chi^{\mathrm{rel}}(\omega)$. Demanding causality for the response function, ($\chi^{\mathrm{rel}}(\tau) = 0$ for $\tau > 0$) we can transform it to the frequency domain as

$$\mathbf{P}(\omega) = \epsilon_r\epsilon_0\chi^{\mathrm{rel}}(\omega)\mathbf{E}(\omega). \tag{2.4}$$

A small change in the relative susceptibility will affect the complex refractive index as

$$\delta\tilde{n} = \sqrt{(1+\chi^{\mathrm{rel}}(\omega))\epsilon_r} - \sqrt{\epsilon_r} \approx \sqrt{\epsilon_r}\chi^{\mathrm{rel}}(\omega)/2, \tag{2.5}$$

where we can relate the change in absorption to the imaginary part as

$$\alpha \approx \frac{\omega\sqrt{\epsilon_r}}{c}\Im\{\chi^{\mathrm{rel}}\}. \tag{2.6}$$

Similarly, the real part of the susceptibility affects the refractive index of the material. For lasers, the negative absorption, or *gain*, is the central quantity needed to quantify the possible enhancement of an ac field through stimulated emission. From the above discussion we learn that for understanding how a material responds to radiation, the susceptibility is the key quantity. In the following we will calculate it through a perturbative approach, which will then be compared to the treatment in our full model described in Chap. 3.

2.2 Susceptibility through time-dependent perturbation theory

Above, the polarization was related to the abstract quantities of susceptibility and the dielectric function, but what is giving rise to the polarization on the microscopic level? We can write the polarization density as a sum over all dipoles N in the material that we study, normalized over the volume V as

$$\mathbf{P}(t) = \frac{1}{V}\sum_i -e\langle\hat{\mathbf{z}}_i(t)\rangle = -e\frac{N}{V}\langle z(t)\rangle\hat{\mathbf{e}}_z \tag{2.7}$$

which we interpret as the density $n = N/V$ times the mean value of the dipole operator. In the above expression we have written the dipole operator $-ez\hat{\mathbf{e}}_z$ in Lorenz (or dipole) gauge. In this work, e is the elementary charge and thus $-e$ is the electron charge. As stated above, we neglect what happens in the background of lattice atoms, and study instead the densities of electrons in the envelope functions of the quantum wells. Our task is thus to calculate how the expectation value of the quantum mechanical operator $z(t)$ changes over time, in the presence of an electro-magnetic field. In the Interaction picture this is defined as

$$\langle z(t)\rangle = \langle\Psi(t)|z(t)|\Psi(t)\rangle \tag{2.8}$$

where $\Psi(t)$ describes the state of our system. Here we will treat the electromagnetic field as a perturbation, switched on at t_0 in a system that was previously

in equilibrium. The system will undergo changes in order to respond to this perturbation, and the refractive index and the absorption will change, as discussed above. For small perturbations, i.e. weak fields, it is often enough to calculate the response to first order in the field strength, why this is called *linear response*. If the fields are stronger, one has to go beyond this approximation and include also the *non-linear* terms in the perturbative expansion. In the following we will outline how the non-linear susceptibility can be calculated perturbatively to high orders in the field strength. This approach will provide insight into the results of the research papers, and give an introduction to non-linear optical response of materials.

In the Interaction picture, the time evolution of the operators is governed by the time-independent part of the Hamiltonian H_o, while the states are propagated by the perturbative, time-dependent $V(t)$. We have $z(t) = \mathrm{e}^{\mathrm{i}H_0 t} z \, \mathrm{e}^{-\mathrm{i}H_0 t}$ and the time evolution operator connecting states at different times:

$$\mathcal{U}(t, t_0)|\Psi(t_0)\rangle = |\Psi(t)\rangle \tag{2.9}$$

where all operators should be interpreted in the Interaction picture. Applying the Schrödinger equation to this equality, we arrive at the operator equation for the time evolution operator

$$\frac{\partial \mathcal{U}(t, t_0)}{\partial t} = V(t)\mathcal{U}(t, t_0) \tag{2.10}$$

$$\mathcal{U}(t, t_0) = -\frac{\mathrm{i}}{\hbar} \int_{t_0}^{t} \mathrm{d}t' V(t')\mathcal{U}(t', t_0) \tag{2.11}$$

where the second line is obtained using the boundary condition $\mathcal{U}(t_0, t_0) = 1$. An iterative solution can be written down as

$$\begin{aligned} \mathcal{U}(t, t_0) = 1 &- \frac{\mathrm{i}}{\hbar} \int_{t_0}^{t} \mathrm{d}t' V(t') \\ &- \frac{1}{\hbar^2} \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t'' V(t') V(t'') + \cdots \end{aligned} \tag{2.12}$$

which lays the foundation for a systematic perturbative expansion, up to all orders in the interaction Hamiltonian. The zero order provides the unperturbed system, the first order can couple two states, and at third order, more complicated situations become possible.

Let us calculate the response of the dipole operator to first order in the interaction. Inserting the first two terms of our expansion into Eq. (2.8) we can calculate

$$\begin{aligned}\delta\langle z(t)\rangle =&\langle\Psi(t)|z(t)|\Psi(t)\rangle - \langle\Psi(t_0)|z(t)|\Psi(t_0)\rangle\\ =&\langle\Psi(t_0)|\mathcal{U}^\dagger(t,t_0)z(t)\mathcal{U}(t,t_0)|\Psi(t_0)\rangle - \langle\Psi(t_0)|z(t)|\Psi(t_0)\rangle\\ =&\langle\Psi(t_0)|\left(1+\frac{i}{\hbar}\int_{t_0}^{t}V(t')dt'\right)z(t)\left(1-\frac{i}{\hbar}\int_{t_0}^{t}V(t')dt'\right)|\Psi(t_0)\rangle\\ &-\langle\Psi(t_0)|z(t)|\Psi(t_0)\rangle\\ =&\langle\Psi(t_0)|\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\left[V(t'),z(t)\right]|\Psi(t_0)\rangle+\mathcal{O}(V^2(t))\end{aligned} \tag{2.13}$$

where only terms linear in the interaction has been kept and terms of higher order are contained in $\mathcal{O}(V^2(t))$. The interaction Hamiltonian for a classical electromagnetic field is given in Paper II by $V(t) = eF(t)z$, where $F(t)$ is the $\hat{\mathbf{z}}$ component of Eq. (2.1) and does not involve any operators. Plugging in the interaction Hamiltonian, we find an expression for the polarization density, following Eq. (2.7), as

$$P_z(t) = -en\delta\langle z(t)\rangle = -\mathrm{i}\frac{e^2}{\hbar}\int_{t_0}^{t}\mathrm{d}t'\langle\Psi(t_0)|\left[z(t'),z(t)\right]|\Psi(t_0)\rangle F(t'). \tag{2.14}$$

In this expression now identify our first order relative susceptibility from Eq. (2.4) as

$$\chi^{(1)}(t-t') = -\mathrm{i}\frac{e^2 n}{\epsilon_0\epsilon_r\hbar}.\langle\Psi(t_0)|\left[z(t'),z(t)\right]|\Psi(t_0)\rangle \tag{2.15}$$

The next step will be to evaluate the expectation value of the commutator using the density matrix of the unperturbed system in the infinite past $t_0 = -\infty$,

$$\langle A\rangle = \sum_k p_k\langle\phi_k|A|\phi_k\rangle \tag{2.16}$$

where p_k denotes the probability of finding the state k in a measurement. We evaluate the matrix elements in this eigenbasis of the unperturbed system of H_0 as

$$\begin{aligned}\langle\alpha|z(t)|\beta\rangle &= \langle\alpha|\mathrm{e}^{\mathrm{i}H_0(t-t_0)}z\mathrm{e}^{-\mathrm{i}H_0(t-t_0)}|\beta\rangle\\ &= \sum_{\gamma,\nu}\langle\alpha|\mathrm{e}^{\mathrm{i}H_0(t-t_0)}|\gamma\rangle\langle\gamma|z|\nu\rangle\langle\nu|\mathrm{e}^{-\mathrm{i}H_0(t-t_0)}|\beta\rangle\\ &= \sum_{\gamma,\delta}\mathrm{e}^{\mathrm{i}E_\alpha(t-t_0)/\hbar}\delta_{\alpha\gamma}\langle\gamma|z|\delta\rangle\mathrm{e}^{-\mathrm{i}E_\nu(t-t_0)/\hbar}\delta_{\nu\beta}\\ &= \langle\alpha|z|\beta\rangle\mathrm{e}^{\mathrm{i}\omega_{\alpha\beta}(t-t_0)} = z_{\alpha\beta}\mathrm{e}^{\mathrm{i}\omega_{\alpha\beta}(t-t_0)}\end{aligned} \tag{2.17}$$

where $\omega_{\alpha\beta} = (E_\alpha - E_\beta)/\hbar$, and we used the fact that H_0 is incapable of coupling different eigenstates. We now introduce the relative time coordinate $\tau = t - t'$ (always positive due to causality) and introduce a protection factor of $\mathrm{e}^{-\eta\tau}$ that will allow us to gradually switch on the interaction as we let $t_0 \to -\infty$ as in Eq. (2.3). At the end of the calculation we will take the limit $\eta \to 0$. This *adiabatic switching* on of the interaction avoids the problem of a sudden interaction part of the Hamiltonian appearing for example at $t = 0$ [44]. Using the density operator of Eq. (2.16) and inserting a complete set of states we arrive at

$$\begin{aligned}\chi^{(1)}(\tau) &= -\frac{\mathrm{i}e^2 n}{\epsilon_0\epsilon_r\hbar}\sum_\alpha p_\alpha \sum_\beta \left[\langle\alpha|z|\beta\rangle\langle\beta|z(\tau)|\alpha\rangle - \langle\alpha|z(\tau)|\beta\rangle\langle\beta|z|\alpha\rangle\right]\mathrm{e}^{-\eta\tau} \\ &= -\frac{\mathrm{i}e^2}{\epsilon_0\epsilon_r\hbar}\sum_\alpha n_\alpha \sum_\beta \left[z_{\alpha\beta}z_{\beta\alpha}\mathrm{e}^{-\mathrm{i}\omega_{\alpha\beta}\tau} - z_{\alpha\beta}z_{\beta\alpha}\mathrm{e}^{\mathrm{i}\omega_{\alpha\beta}\tau}\right]\mathrm{e}^{-\eta\tau}, \quad (2.18)\end{aligned}$$

which brings us directly to the frequency space expression

$$\begin{aligned}\chi^{(1)}(-\omega;\omega) &= \frac{\mathrm{i}e^2}{\epsilon_0\epsilon_r\hbar}\sum_\alpha n_\alpha \sum_\beta \left[\frac{|z_{\alpha\beta}|^2}{\mathrm{i}(\omega - \omega_{\alpha\beta} + \mathrm{i}\eta)} - \frac{|z_{\alpha\beta}|^2}{\mathrm{i}(\omega + \omega_{\alpha\beta} + \mathrm{i}\eta)}\right] \\ &= -\frac{e^2}{\epsilon_0\epsilon_r\hbar}\sum_\alpha n_\alpha \sum_\beta \left[\frac{|z_{\alpha\beta}|^2}{\omega_{\alpha\beta} - \omega - \mathrm{i}\eta} + \frac{|z_{\alpha\beta}|^2}{\omega_{\alpha\beta} + \omega + \mathrm{i}\eta)}\right], \quad (2.19)\end{aligned}$$

where we have used the conventional way of writing the frequency argument of χ. It is read as "the χ coupling an electric field at ω to the polarization of ω", which might be trivial in this case, but becomes practical for the higher order terms (cmf. Eq. (2.23)). The frequencies separated by the semicolon should add up to zero. The poles of Eq. (2.19) are in the lower complex half plane, as we expect for a *retarded* function. As we will use this concept frequently in the thesis, we shall here give it a short introduction. Retardedness means that only events in the past has any meaning to an object. Mathematically we can see it from the integral representation of the Heaviside function

$$\theta(\tau) = -\frac{1}{2\pi\mathrm{i}}\int \frac{1}{\omega + \mathrm{i}\eta}\mathrm{e}^{-\mathrm{i}\omega\tau} \quad (2.20)$$

where η is a small positive number. This has a similar structure as the susceptibility of Eq. (2.19), with a pole at $\omega = -\mathrm{i}\eta$ in the lower complex plane. For $\tau < 0$

the integral has to be closed in the upper half plane, which does not include the pole and the function is thus zero. For $\tau > 0$ on the other hand, the integral has to be closed in the lower half plane and the function is equal to one. Functions with a retarded pole structure in frequency space will carry a Heaviside function when Fourier transformed back to the time domain. This will ensure the correct time dependence in Eq. (2.3) when going from the frequency to the time domain. Negative frequencies in the susceptibility brings the poles to the upper complex plane and gives the function an *advanced* time behavior, the opposite to retardation. This is also true for the complex conjugate, and we will relate this to the Green's functions in Chap. 3.

Let us now evaluate the complex part of $\chi^{(1)}(-\omega;\omega)$ for a system described by states α and β. As we saw above in Eq. (2.6), this will give us the gain coefficient. Using the Dirac identity and considering only positive frequencies, we can pair the terms of the double sum as

$$\Im\{\chi^{(1)}(-\omega;\omega)\} = -\frac{e^2}{\epsilon_0\epsilon_r\hbar}\sum_{\alpha>\beta}(n_\alpha - n_\beta)|z_{\alpha\beta}|^2\pi\delta(\omega_{\alpha\beta}-\omega) \tag{2.21}$$

where we see that we the sign of each term depends on the relative arrangement of populations $n_\alpha - n_\beta$, where we have assumed $E_\alpha > E_\beta$. If we now, following Ref. [45], assume a more physical line shape such as a Lorentzian with a full width half maximum (FWHM) Γ, we find the gain, or negative absorption, in first order of the field strength as

$$\begin{aligned} G(\omega) &\approx -\frac{\omega\sqrt{\epsilon_r}}{c}\Im\{\chi^{(1)}(-\omega;\omega)\} \\ &= \frac{\omega\sqrt{\epsilon_r}}{c}\frac{e^2}{\epsilon_0\epsilon_r}\sum_{\alpha>\beta}(n_\alpha - n_\beta)|z_{\alpha\beta}|^2\pi\frac{1}{2\pi}\frac{\Gamma}{\hbar^2(\omega_{\alpha\beta}-\omega)^2+\Gamma^2/4} \\ &= \frac{\hbar\omega e^2}{2c\epsilon_0\sqrt{\epsilon_r}}\sum_{\alpha>\beta}(n_\alpha - n_\beta)|z_{\alpha\beta}|^2\frac{\Gamma}{\hbar^2(\omega_{\alpha\beta}-\omega)^2+\Gamma^2/4}, \end{aligned} \tag{2.22}$$

which becomes particularly simple for a 2-level system. This expression is very useful if we want to understand how to design a device with a large gain coefficient. Mainly three things are important, the dipole matrix element, the inversion and the width of the transition Γ. Although it is only to first order in the field strength, it is a very strong result if one also has information on how the inversion decays at higher field strengths. This expression is used near resonance in Paper V to study

the effects of a changing scattering environment on the dipole matrix elements $z_{\alpha\beta}$.

The above formula can be directly calculated from Fermi's Golden Rule, for example as outlined in Ref. [45], or in textbooks on quantum mechanics such as Ref. [44], however the formalism used here allows us to extend the calculations to next order easily. Including higher terms in the expansion of the time evolution operator, one can isolate terms of the same order in the interaction. Here, we will display the end result for the second order as it is used to approximate the second harmonic generation in Paper III. To second order we have

$$\chi^{(2)}(-\omega;\omega_1,\omega_2) = \frac{e^3}{\epsilon_0\epsilon_r\hbar^2}\sum_\alpha n_\alpha \sum_{\beta,\gamma}\left\{\frac{z_{\alpha\beta}z_{\beta\gamma}z_{\gamma\alpha}}{(\omega_{\alpha\gamma}+\omega_2)(\omega_{\alpha\beta}+\omega)}\right. \tag{2.23}$$
$$\left. - \frac{z_{\alpha\beta}z_{\beta\gamma}z_{\gamma\alpha}}{(\omega_{\alpha\gamma}+\omega_2)(\omega_{\beta\gamma}+\omega)} - \frac{z_{\alpha\beta}z_{\beta\gamma}z_{\gamma\alpha}}{(\omega_{\beta\alpha}+\omega_2)(\omega_{\beta\gamma}+\omega)} + \frac{z_{\alpha\beta}z_{\beta\gamma}z_{\gamma\alpha}}{(\omega_{\beta\alpha}+\omega_2)(\omega_{\gamma\alpha}+\omega)}\right\}$$

where we see that the dipole matrix element enters in powers of three. For second harmonic generation, the frequencies ω_1 and ω_2 are equal and the terms can be paired up, similar to what was done in the first order calculation, to form the expression in Paper III. Here, we have defined the polarization including ϵ_r in Eq. (2.3), to be consistent with the calculation of gain. In Paper III however, the definition did not include this which explains the extra ϵ_r in the denominator above (for polarization it does not matter as it cancels out with the susceptibility).

We will now write out the polarization in the series expansion and collect the results so far. To simplify the expression, and to adapt it to the content of this thesis, we will include only one time dependent electric field at ω_0 in our calculations. Neglecting the spatial dependence, we have schematically

$$\begin{aligned}\mathbf{P}(\omega) =& \epsilon_0\epsilon_r\chi^{(1)}(-\omega;\omega_0)\mathbf{E}(\omega_0)\\ &+\epsilon_0\epsilon_r\chi^{(2)}(-\omega;\omega_0\pm\omega_0)(\mathbf{E}(\omega_0))^2\\ &+\epsilon_0\epsilon_r\chi^{(3)}(-\omega;\omega_0\pm\omega_0\pm\omega_0)(\mathbf{E}(\omega_0))^3+\cdots\end{aligned} \tag{2.24}$$

where different physical phenomena now can be singled out easily; at first order, or linear response as it is usually denoted, we have absorption and gain as defined above. At this order the dipole matrix element enters as a square as shown in Eq. (2.22). All terms beyond the first order are what is commonly referred to as *non-linear* optical processes. At second order we have sum frequency generation

or second harmonic generation, as well as difference frequency generation. Here we can also include processes that are of non-linear nature but not necessarily frequency dependent, as the arguments can cancel each other. As shown in Eq. (2.23) the dipole matrix element enters these expressions in powers of three.

In the third order susceptibility, three frequencies can interact to influence a fourth one. This is known as four-wave mixing and it is the physical effect responsible for mode locking and the generation of frequency combs, which is a rather new achievement in the field of QCLs [46, 38]. At this order, we also have the Kerr effect, which describes the change of refractive index in a material under irradiation [2]. The simultaneous change in the imaginary part is consequently the gain saturation, which is central in the work presented in this thesis. As we only consider one field, it can be seen from the expansion in Eq. (2.24) that the field at frequency ω_0 can couple back to itself. Although the complete calculation is too lengthy to fit into this thesis, complete expressions can found in Ref. [47]. In Chap. 5 we will provide a quicker way of estimating gain saturation and show that the dipole matrix enters to the fourth power, as expected by the lower order terms.

2.3 Our approach to the susceptibility

We will conclude this chapter by discussing briefly how we approach the gain concept without a perturbative expansion in the field strength. In our simulations we calculate the frequency dependent current which we can relate to the conductivity as $\mathbf{J}(\omega) = \sigma(\omega)\mathbf{E}(\omega) = -\mathrm{i}\omega\mathbf{P}(\omega)$. Relating to the susceptibility in Eq. (2.4) we find that $\chi(\omega) = \mathrm{i}\sigma(\omega)/\omega\epsilon_0\epsilon_r$ which allows us to express the gain as

$$G(\omega) = -\frac{\omega n_r}{c}\Im\{\chi^{\mathrm{rel}}(\omega)\} = -\frac{1}{c\epsilon_0\sqrt{\epsilon_r}}\Re\{\sigma(\omega)\}\,. \tag{2.25}$$

The conductivity can be extracted from the frequency dependent current. Instead of a perturbation expansion in the ac field strength, we expand our observables in frequency, following Ref. [48], to be able to go to higher intensities. This allows us to calculate the different current components $J_0, J_1(\omega_0), J_2(2\omega_0), \ldots$ which is truncated at some chosen maximum harmonic. The current, or conductivity, is thus expanded as

$$J(t) = \sum_h J_h \mathrm{e}^{\mathrm{i}h\omega_0 t}, \tag{2.26}$$

up to a certain $h_{\max}$. Compared to an expansion in field strength, this expansion will contain parts of all orders in the standard perturbation expansion, but will be limited to the frequency of the laser field ω_0 and its higher harmonics. By this truncation of the frequency space, the approximation enters instead as a restriction of the phase space of the system. The system is allowed to oscillate only at the frequencies included. In this procedure, indirect couplings via higher harmonic orders are cut off. In Sec. 3 we provide the details of how this is implemented how the Fourier components of different frequencies interact.

To see the effects that we capture with a single pumping field more clearly, we insert the perturbative expansion for the polarization into the exact gain expression above in Eq. (2.25). Only the odd terms can couple back to the same frequency as the pumping field, which leads to

$$G(\omega) = -\frac{\omega n_r}{c}\left[\chi^{(1)}(\omega) + \chi^{(3)}(\omega)E^2(\omega) + \chi^{(5)}(\omega)E^4(\omega)\cdots\right] \tag{2.27}$$

where it is possible to identify the linear gain term as well as the saturation which enters with the square of the field. The second order susceptibility couples to the second harmonic, which is discussed in Chap. 3 and Chap. 5, but also to the zero frequency. For gain with $\omega = 0$, only even terms enter to give

$$G(0) = -\frac{\omega n_r}{c}\left[\chi^{(1)}(0) + \chi^{(2)}(0)E(\omega) + \chi^{(4)}(0)E^3(\omega)\cdots\right], \tag{2.28}$$

which shows that high intensity terms enter in the non-linear orders. The gain was linked to the conductivity in Eq. (2.25). This demonstrates how the static conductance can change as we increase the ac field strength. The phenomena is known as *dynamical localization* and has been shown to decrease the tunneling current through SLs [49, 50, 51]. In a centro-symmetric media, all even terms of the susceptibility are zero, so in order to observe this effect, an external field has to be applied to break the symmetry of the SL. In optics, the static contribution of these terms are referred to as *optical rectification* [2].

3 Non-equilibrium Green's functions

In this chapter we outline our method of choice for modeling the heterostructures described in Chap. 1, and show how we apply it to get quantitative results for quantum cascade lasers under operation. We will begin by motivating why this is the best trade-off between accuracy and computational time, and then we will discuss a few, although not all, details of our implementation. The theoretical background in Sec. 3.2 will closely follow the discussion in Paper II, which outlined the latest implementation of the model at the present time. After this we will discuss new contributions that are part of this thesis. One example is in Sec. 3.5, where we discuss how we can treat electron-electron scattering in the single plasmon pole approximation (SPPA), effectively introducing Paper V.

3.1 Motivation

As discussed in Chap. 2, the density matrix is the key quantity to calculate. From this, we can extract the current and conductivity. The time evolution of the density matrix is governed by the model Hamiltonian for a QCL system

$$H(t) = H_{\mathrm{SL}} + H_{\mathrm{MF}}(t) + H_{\mathrm{ac}}(t) + H_{\mathrm{scatt}} \tag{3.1}$$

where H_{SL} is the superlattice part containing the conduction band potential and the applied bias over the structure, $H_{\mathrm{MF}}(t)$ is the meanfield which contains both a static and a time dependent part and $H_{\mathrm{ac}}(t)$ includes a classical oscillating elec-

tromagnetic field in Lorenz gauge [11]

$$H_{\mathrm{ac}}(t) = eF_{\mathrm{ac}} z \cos(\omega_0 t) \tag{3.2}$$

where we will will often refer to the ac field strength $eF_{\mathrm{ac}}d$ with d being the length of one period, in order to have units of eV. The last term H_{scatt} contains scattering between states of different $\mathbf{k}$, representing the degrees of freedom in the lateral directions. It is the only part of Eq. (3.1) that breaks the translational invariance and is thus, in contrast to the other terms, not diagonal in $\mathbf{k}$. We will treat this perturbatively using a similar expansion as done in Chap. 2, but expanded in orders of the scattering potential instead.

It turns out that solving the equation of motion for the density matrix subject to the Hamiltonian of Eq. (3.1) is a non-trivial task for THz QCLs due to the energy scales of the system. As the tunneling, scattering, and emission of photons are all of the same order in energy, general simplifications that are also accurate are hard to find. In this section we provide a short overview of the approaches used in different groups. For a helpful review on the topic please see Ref. [36].

The simplest approach, which also provides a lot of physical insight is to calculate the intersubband scattering rates for thermalized electron subband distributions [52, 53, 54]. Then one can solve these rate equations together with phenomenological tunneling amplitudes, and then calculate the intersubband gain with the linear response formula, or Fermi's golden rule, in Eq. (2.22). As the method is straightforward, it is possible to include many scattering mechanisms, and this approach was used during the early stages of QCL research to study hot phonon effects [55, 56], confined phonons [57] as well as electron-electron (e-e) scattering [58, 59]. Monte Carlo methods have been used to calculate the $\mathbf{k}$-dependent scattering rates between the non-equilibrium electron subband distributions [60, 61]. One strength of this method is that e-e scattering can be taken into account at a small computational cost [62, 63, 64, 65] which is harder in more complex methods.

The standard perturbative schemes for density matrix approaches solves the von Neumann equation

$$\mathrm{i}\hbar \frac{\partial \rho}{\partial t} = [H, \rho] \tag{3.3}$$

either by propagating in time to a stationary solution or directly, by matrix inversion for stationary solutions [66, 67, 68, 69, 70]. This has also been performed

with a **k**-resolved density matrix [62, 71, 72]. One major problem with these approaches, is that negative densities is a general feature of the way that the perturbation series is treated, as discussed in Ref. [72].

As the semi-classical methods of rate equations and Monte-Carlo by definition will not be able to describe fully the quantum aspects of charge transport trough QCLs and the density matrix methods provide unphysical results, it is both interesting and necessary to do something more advanced. One picture that is helpful when we want to understand the nature of charge transport is the tunneling across one barrier. Once the system is out of equilibrium, the states are broadened in energy; that is, each state $|\alpha, \mathbf{k}\rangle$ now has its own energy dispersion. Even if two states are not completely aligned, the upper tail of one state might overlap with the lower tail of another, and vice versa. To proper handle such situations, one either has to include scattering assisted tunneling in the density matrix approach [70] or introduce a second time into the equation of motion like Eq. (3.3). The latter approach supplies a memory function into the system, which is very valuable. If one can excite at one time, and probe at another, we can resolve scattering times and state lifetimes. This is accomplished by the NEGF formalism which is now used by several groups for QCL simulations [73, 41, 74, 75, 76]. Together with the correct energy broadening of the states and coherences, the formalism includes scattering correlations into the gain linewidth, shown to be of importance for interface roughness [77, 78].

3.2 Theoretical background

We will in the following use *occupation number* or *second quantization* formalism, where the field operators $a_\alpha^\dagger, a_\beta$ are used to create or annihilate particles, in states α and β, respectively.

The physics of the SL can be described by the Bloch states

$$\Psi_\alpha(\mathbf{r}) = \mathrm{e}^{\mathrm{i}\mathbf{k}_\alpha \cdot \mathbf{r}} u_\alpha(\mathbf{r}) \tag{3.4}$$

enumerated by α for ground state, excited state and so on, where $u_\alpha(\mathbf{r})$ is a lattice periodic function. These envelope functions are simultaneous eigenfunctions to the SL Hamiltonian and to the translation operator, taking us from one lattice site, or in our case one SL period, to another. As these states are delocalized in

nature, we use instead a localized superposition of them as a basis for calculations. This type of states are called Wannier states [79] and are practical for calculations as they are invariant under an applied potential. The Wannier states were plotted in Fig. 1.4 for the SL in Chap. 1. These states will often be referred to as *subbands*, with their discrete energy being at the center of the respective miniband [11]. There is a certain freedom in the construction of these states; here we follow Ref. [80] to create maximally localized functions, as this simplifies the calculation of matrix elements. We use an effective two band model to take non-parabolicity effects into account, as detailed in the thesis of Franckié [81]. The two-band model was not used in Paper I or II of this thesis.

When we apply the NEGF method to the transport problem in QCLs, the main object of interest is the lesser Green's function

$$G^{<}_{\alpha\beta}(\mathbf{k}; t_1, t_2) = \mathrm{i}\langle a^{\dagger}_{\beta}(t_2) a_{\alpha}(t_1)\rangle \tag{3.5}$$

which bears close resemblance to the density matrix. In fact the lesser Green's function is defined here to match the density matrix at equal times, that is

$$G^{<}_{\alpha\beta}(\mathbf{k}; t_1, t_1^{+}) = \mathrm{i}\rho_{\beta\alpha}(\mathbf{k}; t_1) \tag{3.6}$$

where the "+" indicates that the limit is to be taken from above. The scheme is thus to solve not the wavefunction or the density matrix of the problem, but the Green's function, and then extract the quantities that we are interested in from this object. This way we hope to capture most of the quantum aspects of charge transport and optical response. The resemblance of the lesser Green's function to the density matrix also captures the essence of the non-equilibrium formalism; in equilibrium theory there is no need to evaluate the $\mathbf{k}$ dependence of the density matrix as it will follow Fermi-Dirac statistics.

To complement the lesser Green's function there is also the *retarded* and *advanced* functions defined as

$$\begin{aligned} G^{\mathrm{r}}_{\alpha\beta}(\mathbf{k}; t_1, t_2) &= -\mathrm{i}\left\langle \left\{ a_{\alpha}(t_1), a^{\dagger}_{\beta}(t_2) \right\} \right\rangle \theta(t_1 - t_2) \\ G^{\mathrm{a}}_{\alpha\beta}(\mathbf{k}; t_1, t_2) &= \quad \mathrm{i}\left\langle \left\{ a_{\alpha}(t_1), a^{\dagger}_{\beta}(t_2) \right\} \right\rangle \theta(t_2 - t_1) \end{aligned} \tag{3.7}$$

which live in two different part of time, the past and the future, respectively. Here we have used the heavy-side function ($\theta(t) = 0$ if $t < 0$ and $\theta(t) = 1$ if $t > 0$) as well as the anti commutator ($\{a, b\} = ab + ba$) for the fermionic creation

and annihilation operators considered here. As with the susceptibility, the complex conjugate of the retarded Green's function is the advanced Green's function. These three functions, the lesser, retarded and advanced, can be combined in different linear combinations, and the definition presented here is far from the only one [82]. The most important detail of these special functions is that their time arguments are complex in general. As easy as this is for development of the theory, as complicated it is to bring the final expressions back to the real time axis. This can be done by analytic continuation, and the Langreth rules [83] provide a convenient set of rules to accomplish this.

As detailed in Paper II, the equation of motion for the Green's function is closed using self-energies $\Sigma_{\alpha\beta}(\mathbf{k}; t, t')$ which are functionals of the Green's functions themselves. Here we present the *Dyson integral equation*

$$\begin{aligned}\mathbf{G}^{\mathrm{r/a}}(\mathbf{k}; t_1, t_2) &= \mathbf{G}^{0,\mathrm{r/a}}(\mathbf{k}; t_1, t_2) \\ &+ \int \frac{\mathrm{d}t}{\hbar} \int \frac{\mathrm{d}t'}{\hbar} \mathbf{G}^{0,\mathrm{r/a}}(\mathbf{k}; t_1, t) \mathbf{\Sigma}^{\mathrm{r/a}}(\mathbf{k}; t, t') \mathbf{G}^{\mathrm{r/a}}(\mathbf{k}; t', t_2),\end{aligned} \tag{3.8}$$

where we introduce bold capital notation for matrices. Here we use the bare Green's function $\mathbf{G}^{0,\mathrm{r/a}}$ which we would get from the unperturbed system [85]. This expression can be transformed to a integro-differential equation (Eq. (10) of Paper II) by multiplying from the left with the inverse of $\mathbf{G}^{0,\mathrm{r/a}}$ and performing one of the time integrals. In non-equilibrium theory we also have to solve the lesser Green's function using the Keldysh relation [85]

$$\mathbf{G}^{<}(\mathbf{k}; t_1, t_2) = \int \frac{\mathrm{d}t}{\hbar} \int \frac{\mathrm{d}t'}{\hbar} \mathbf{G}^{\mathrm{r}}(\mathbf{k}; t_1, t) \mathbf{\Sigma}^{<}(\mathbf{k}; t, t') \mathbf{G}^{\mathrm{a}}(\mathbf{k}; t', t_2), \tag{3.9}$$

once the retarded and advanced parts are known. This implies an iterative process where the new Green's functions are used to update the self-energies. As the equations are well described in Paper II, where we also clearly show how to go from the

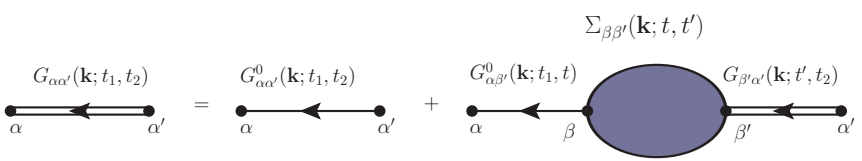

Figure 3.1: Graphical form of the Dyson equation where the blue object labeled as Σ is the self-energy. The propagators are drawn traveling to the left in order to make them directly equivalent to the mathematical definitions in Eq. (3.7). The diagrams shown in this thesis were created using the open source software JaxoDraw [84].

time to the energy domain, we will here instead give a graphical interpretation of the Dyson equation and the self-energies.

The diagram in Fig. 3.1 is the equivalent of the Dyson equation (3.8), and shows how the dressed Green's function (double lines) is built up as combinations of the bare Green's function (single lines) and self-energies (blue blob). Here we refer to *dressed* and *bare* functions instead of *interacting* and *non-interacting* as we most often neglect the Coulomb interaction beyond meanfield effects. The one-particle propagators are instead broadened by the scattering events, why we call the iterative solution a *dressed* function.

In the presence of a time dependent ac field, the objects ($\mathbf{G}$'s and $\mathbf{\Sigma}$'s) depend both on the time difference $t_1 - t_2$, and the phase and frequency of the ac field. To incorporate this, we define the Fourier transform as

$$F(\mathbf{k}; t_1, t_2) = \int \frac{\mathrm{d}E}{2\pi} \sum_h \mathrm{e}^{-iE(t_1-t_2)/\hbar} F_h(\mathbf{k}, E) \mathrm{e}^{-\mathrm{i}h\omega_0 t_1} \tag{3.10}$$

where ω_0 is the frequency of the ac field, and the object in energy space now effectively carries an additional energy label h. The system observables are now

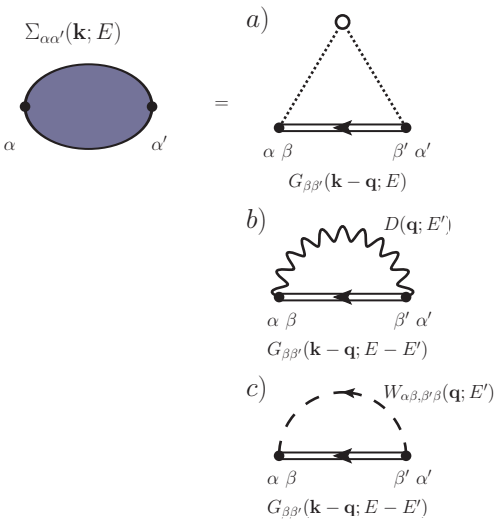

Figure 3.2: Diagrams in energy domain for the three types of physical processes studied in this thesis. They are drawn here in the self-consistent Born-approximation, with elastic scattering in a), phonon scattering in b) connecting the fermionic to the bosonic propagator $D(\mathbf{q}, E')$ as a wiggly line, and the screened e-e interaction mediated through $W_{\alpha\beta\beta'\alpha'}$ in c) drawn as dashed line.

expressed as Fourier series with the fundamental frequency ω_0 and its higher harmonics.

In energy space we keep the same diagrams although the rules to transform them into mathematical expressions are slightly changed. Below we give an example of this procedure for the LO phonon self-energy. The different types of self-energies that we are interested in are shown in Fig. 3.2. In a), the elastic scattering self-energies are shown. These are due to ionized impurities, alloy disorder and interface roughness. The scattering potentials are treated as random distributions, over which we perform a statistical average. The elastic processes connect Green's functions of different $\mathbf{k}$-values and same energy. In contrast, the phonon self-energy in b) does not. Here, a scattering event excites a phonon of energy E' and thus the self-energy couples Green's functions of different energy, making it an inelastic process. The third type drawn in c) demonstrates the screened Coulomb interaction $W_{\alpha\beta\beta'\alpha'}$ which is a scattering event between two electrons interchanging both $\mathbf{q}$ and E'.

3.3 Implementation

As an example of how the method is implemented in our case, we will provide an outline of how the optical phonon self-energy is formulated in the self-consistent Born approximation. To demonstrate the power of the diagrammatic approach, we will formulate the self-energy directly from Fig. 3.2 b). This will be instructive as the procedure for the SPPA self-energy for e-e scattering will follow along the same lines, as discussed in Sec. 3.5. To find the coupling strength $\lambda(\vec{q})$ for the diagram here in Fig. 3.3 we have from Paper II we have the perturbation potential

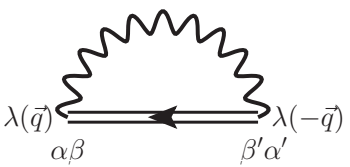

Figure 3.3: Graphical representation of the phonon scattering self-energy in the self-consistent Born approximation. The bosonic (phonon) Green's function (wiggly line) couples to the fermionic one with coupling $\lambda(\vec{q})$.

for LO phonons

$$V_{\text{opt phon}}(\vec{r},t) = \sum_{\mathbf{q}} g(\vec{q}) \mathrm{e}^{\mathrm{i}\vec{q}\cdot\vec{r}} \left[b_{LO}(\vec{q}) + b^{\dagger}_{LO}(-\vec{q}) \right]$$

with the Frölich coupling $g(\vec{q})$, where we will split the wave vector $\vec{q} = \mathbf{q} + q_z \hat{\mathbf{e}}_z$ in the following to take the layered QCL structure into account. From this potential we deduce the coupling of

$$\lambda(\vec{q})_{\alpha\beta} = g(\vec{q}) \int \mathrm{d}z \, \Psi^{*}_{\alpha}(z) \mathrm{e}^{\mathrm{i}q_z z} \Psi_{\beta}(z) \tag{3.11}$$

for the envelope states $\Psi_{\alpha}(z)$ and $\Psi_{\beta}(z)$. This reflects the spatial overlap, and it is common in QCL design to try to minimize this between the upper and lower laser state. Following the diagram rules of Ref. [86] we construct the mathematical self energy expression by (1) associating momentum and energy with each line and apply conservation of both quantities at each vertex, (2) including an interaction at each vertex with appropriate matrix elements, (3) integrating/summing over all internal degrees of freedom. The mathematical expression that we need to evaluate then reads as

$$\begin{aligned} \Sigma_{\alpha\alpha'}(\mathbf{k},E) = & \sum_{\beta\beta'} \sum_{q_z\mathbf{q}} \lambda_{\alpha\beta}(\vec{q}) \lambda^{*}_{\beta'\alpha'}(\vec{q}) \\ & \times \int \frac{\mathrm{d}E'}{2\pi} G_{\beta\beta'}(\mathbf{k}-\mathbf{q}, E-E') D(\mathbf{q},E'), \end{aligned} \tag{3.12}$$

where $D(\mathbf{q},E')$ is the phonon Green's function, drawn as a wiggly line in Fig. 3.3, propagating in the same direction as the Green's function, in order to ensure energy and momentum conservation.

To transform our complex time objects to expressions on the real time axis one performs analytic continuation on the diagram. In our case the product structure of the phonon and electron Green's function is $C(\tau,\tau') = G(\tau,\tau')D(\tau,\tau')$ which according to the Langreth rules [83] should be analytically continued as

$$C^{<}(t,t') = G^{<}(t,t')D^{<}(t',t) \tag{3.13}$$

$$C^{\mathrm{r}}(t,t') = G^{<}(t,t')D^{\mathrm{r}}(t',t) + G^{\mathrm{r}}(t,t')D^{<}(t',t) + G^{\mathrm{r}}(t,t')D^{\mathrm{r}}(t',t), \tag{3.14}$$

for the lesser part and retarded part respectively. The lesser and retarded part of the non-interacting phonon propagators are given in Ref. [85] as

$$D_0^<(\mathbf{q}, E) = -2\pi\mathrm{i}[(N_{\mathrm{LO}}+1)\delta(E+\hbar\omega_{\mathrm{LO}}) + N_{\mathrm{LO}}\delta(E-\hbar\omega_{\mathrm{LO}})] \tag{3.15}$$

$$D_0^{\mathrm{r}}(\mathbf{q}, E) = \mathrm{i}\pi\left[\frac{1}{E-\hbar\omega_{\mathrm{LO}}-i\eta} - \frac{1}{E+\hbar\omega_{\mathrm{LO}}-i\eta}\right], \tag{3.16}$$

where $N_{\mathrm{LO}} = (1-\mathrm{e}^{-\hbar\omega/k_{\mathrm{B}}T})^{-1}$ is the occupation of the phonon mode and $\hbar\omega_{\mathrm{LO}}$ is the LO phonon energy which is assumed to be dispersion-less. These real time propagators allows us to complete the expressions and formulate the lesser self-energy as

$$\begin{aligned}\Sigma^<_{\alpha\alpha'}(\mathbf{k}, E) = &-\sum_{\beta\beta'}\sum_{q_z\mathbf{q}} \lambda_{\alpha\beta}(\vec{q})\lambda^*_{\beta'\alpha'}(-\vec{q}) \\ &\times\left[G^<_{\beta\beta'}(\mathbf{k}-\mathbf{q}, E-\hbar\omega_{\mathrm{LO}})N_{\mathrm{LO}} + G^<_{\beta\beta'}(\mathbf{k}-\mathbf{q}, E+\hbar\omega_{\mathrm{LO}})(N_{\mathrm{LO}}+1)\right],\end{aligned} \tag{3.17}$$

where a similar expression for the retarded self-energy can be found using Eq. (3.14). At this stage we perform a central approximation in our evaluation of the self-energies, where we instead of the full momentum dependence of the coupling strength and overlap use a set of typical momentum transfers $E_{\mathrm{typ}}, E_{\mathrm{typ}'}$. This effectively makes the self-energy independent of $\mathbf{k}$ and leaves only the Green's function still under the summation in $\mathbf{q}$, and allows us to write the expression on the form

$$\Sigma^<_{\alpha\alpha'}(E, \mathbf{k}) = \sum_{\beta\beta'} X^{\mathrm{opt}}_{\alpha\alpha'\beta\beta'} \int \mathrm{d}E_{k'}\, G^<_{\beta\beta'}(E_{k'}, E\pm\hbar\omega_{\mathrm{LO}}) \times \left[N_{\mathrm{LO}} + \frac{1}{2} \pm \frac{1}{2}\right]. \tag{3.18}$$

The elastic self-energies take a similar form but without the energy shift in the Green's function argument and the phonon occupation numbers.

The approximation of using typical energies in the evaluation of our self-energies is necessary in order to keep the computational times reasonable. However, it introduces a fit parameter that needs to be handled with care. In order to find a representative value, we compare to scattering rates calculated using thermalized subbands following Boltzmann distributions, for a set of low doped structures and temperatures. We calculate both intra and inter-subband scattering rates,

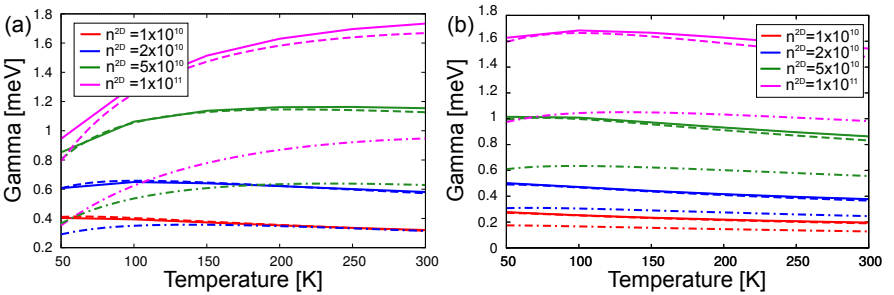

Figure 3.4: Scattering rates as a function of temperature for different parametrizations for the typical energy $E_{\rm typ}$ and different dopings in units of $\rm cm^{-2}$). In (a) and (b), results are presented for the structures of Ref. [87] (46 mV/period) and Ref. [29] (65 mV/period), respectively. The old parametrization (dotted-dashed) was $E_{\rm typ} = 7.4\ \rm meV + k_B T_L$ whereas the best fit yielded $E_{\rm typ} = 0.5\ \rm meV + 0.4 k_B T_e$ in (a) and $E_{\rm typ} = 2.0\ \rm meV + 0.4 k_B T_e$ in (b) (dashed). Solid lines are reference calculations using thermalized subbands.

and compare them to rates extracted from the self-energies when using different values of $E_{\rm typ}$. Part of our results are displayed in Fig. 3.4 (a) and (b), for a four-well [87] and three-well [29] THz QCL, respectively. Inter-subband impurity scattering rates for the strongest transitions in each structure are shown, for four different sheet doping densities. For both structures, we compare the best fit to the old parametrization, reading $E_{\rm typ} = 7.4\ \rm meV + k_B T_L$, with T_L being the lattice temperature. The results vary slightly from structure to structure, although it is clear from Fig. 3.4 that the parameter $E_{\rm typ}$ was too high in the old parametrization, giving too little scattering. Weighing together results also for intra-subband rates, and from interface roughness scattering, the final parametrization reads $E_{\rm typ} = (0.5 - 1.0)\ \rm meV + k_B T_e$, with electron temperature T_e. Assuming that the electron temperature is about 50 K higher than the lattice temperature [88], the parametrization becomes $E_{\rm typ} \approx 3\ \rm meV + k_B T_L$ where we have now replaced T_e with T_L, as the latter is the input parameter to the NEGF simulations.

This parametrization is then kept for all calculations. Papers I-III used the old parametrization, while the new, improved version, was used in Papers IV-VI. As a additional complication, the real part of the retarded version of Eq. (3.18) will have a logarithmic divergence. This is treated by introducing a cutoff in the **k**-space of the integral, as described in Appendix A of Paper VI.

3.4 Frequency expansion in the Green's function

With our definition of the Fourier transform in Eq. (3.10), an additional energy is introduced, given by the h-index of the Green's function. The Green's function matrices of different h-index are not independent, but coupled by the applied optical field of frequency ω_0 and field strength $eF_{\mathrm{ac}}d$. The Dyson Eq. (3.8) becomes a block-matrix equation that we solve by inverting the total tensor to get the Green's functions of all indexes h. In Fig. 3.5 the block matrix structure is shown, and it is

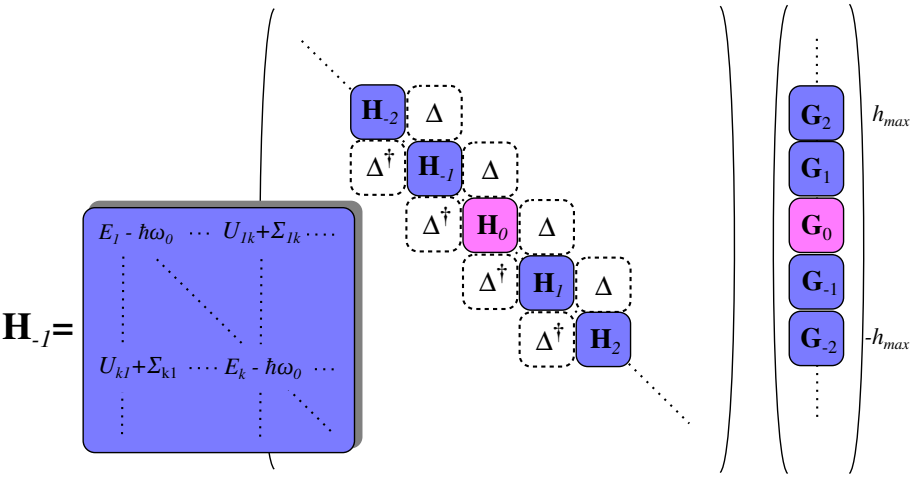

Figure 3.5: Shematic structure of the matrix Dyson equation. The inset shows the structure within each photo-block. Here U_{kl} are the parts of the Hamiltonian of Eq. (3.1) diagonal in $\mathbf{k}$. $\Delta^\dagger$ contains the couplings between each photo-block consisting of the optical field, meanfield and self-energy parts. The elements shown correspond to $h_{\max} = 2$.

the Green's function analogue of the well known Floquet expansion for periodic Hamiltonians [89]. This expansion is exact in the field strength and thus includes the non-linear response terms of Eq. (2.24) to all orders. A truncation is however necessary for practical reasons, but it has been shown by Brandes [48] that a matrix consisting only of the first few orders of photo-blocks is enough to yield close to exact results. When we simulate samples under high intensity, like in Paper III and VI, the index at which we perform the truncation ($h_{\max}$) becomes a parameter that we increase until convergence is reached.

3.5 Electron-electron scattering

In the above implementation and in Paper II, we have seen how the system is solved, both with a superlattice and a scattering Hamiltonian for non-interacting particles. The only way the electrons communicate is via the mean field that enters the superlattice Hamiltonian, and through the Pauli exclusion principle that enters in the formulation of the scattering self-energies. Although this approximation seems reasonable; in Paper VI we see overall good agreement with experiments, we would like to study the relevance of e-e scattering. We believe this to be especially important at low temperatures, where the acoustic phonon scattering is weak and no other inelastic processes are present to thermalize the electron gas. In this section we show how to implement a rudimentary form of the GW equations [90] for QCLs.

Previous studies have given somewhat contradicting results. By analyzing scattering rates Harrison [59] found that e-e scattering provides a quick subband thermalization. In Ref. [62], a Monte-Carlo study using two-body e-e scattering as well as optical phonon scattering, showed that the inclusion of the e-e term was crucial to explain the transport properties of a mid-IR QCL. On the contrary, another Monte-Carlo study of Ref. [63] where elastic scattering was also included, the relative importance of e-e scattering versus electron-impurity scattering was shown to be small, other than for carrier thermalization. In NEGF studies, e-e interaction self-energies have been included on the level of Hartree with static screening [75]. Here, it was found only to play a minor role, as electrons rather thermalized via double-phonon emission over several periods than via e-e scattering. In Paper IV, the aim was to include an effective form of the e-e scattering beyond the Hartee-Fock approximation. These results were later used in Paper VI to estimate the influence of e-e scattering on a large number of devices. Here we will give a short introduction to the difficulties that requires the use of the approximations discussed in Paper IV.

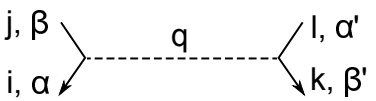

Figure 3.6: Feynman diagram for e-e scattering with time going in the upward direction. The momentum and energy transfer is q and ω respectively.

Screening

The e-e scattering event that we want to study can be described by the diagram of Fig. 3.6. Two electrons exchange a photon that carries the momenta $|\mathbf{q}| = q$. We describe the Coulomb interaction potential as

$$V(|\mathbf{r} - \mathbf{r}'|) = \frac{e^2}{4\pi\epsilon_0\epsilon_r}\frac{1}{|\mathbf{r} - \mathbf{r}'|}, \tag{3.19}$$

which is a long range interaction thanks to the slow algebraic decay. If we consider the standard 3D Fourier transform we have

$$V(\mathbf{q}) = \frac{e^2}{\epsilon_0\epsilon_r}\frac{1}{q^2}. \tag{3.20}$$

The divergence at $q = 0$ is a consequence of the infinite range of the interaction. In solid state materials, the charges are not isolated but part of a dense electron gas. This effectively screens the interaction, which can be understood as a reconfiguration of the orbitals of the other electrons in the material, to counteract the Coulomb forces. Let us consider the screened potential of a point charge, given by

$$V^{\text{eff}}(|\mathbf{r} - \mathbf{r}'|) = \frac{e^2}{4\pi\epsilon_0\epsilon_r}\frac{1}{|\mathbf{r} - \mathbf{r}'|}\mathrm{e}^{-\lambda r}, \tag{3.21}$$

which has an additional exponential decay with a characteristic length of λ. Looking now at this Fourier transform compared to Eq. (3.20), we find

$$V^{\text{eff}}(\mathbf{q}) = \frac{e^2}{\epsilon_0\epsilon_r}\frac{1}{q^2 + \lambda^2} = \frac{V(\mathbf{q})}{\epsilon(\mathbf{q})} \tag{3.22}$$

where the singularity is cured. Here we have introduced the dielectric function $\epsilon(\mathbf{q})$ to describe the screening, and expressed the screened interaction $V^{\text{eff}}(\mathbf{q})$ in terms of $\epsilon(\mathbf{q})$ and the bare interaction $V(\mathbf{q})$. This particular case is called static screening and has no energy dependence. We use it in the impurity scattering self-energies detailed in Paper II, with the characteristic length λ related to an effective electron temperature and density. For more complicated systems evolving in time, as the electron distributions in QCL, the screening is also energy dependent, i.e. *dynamic*, with $\epsilon(\mathbf{q}) \rightarrow \epsilon(\mathbf{q}, E)$ and $V(\mathbf{q}) \rightarrow W(\mathbf{q}, E)$, where W denotes the screened interaction.

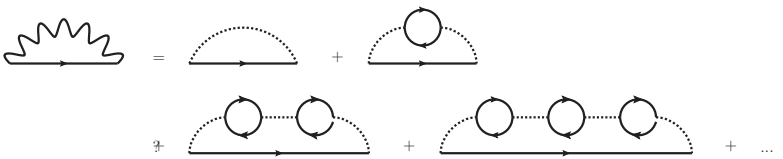

Figure 3.7: The first few diagrams contained in the GW approximation. The interaction is self-consistently screened by the polarization in the medium, represented here as circles.

Calculating the dielectric function

As discussed in Chapter 2, the dielectric function relates the polarization to the electric field. A common approach to calculate the dielectric function, or the screened interaction, is the GW approximation [90]. The name comes from the fact that we use one dressed Green's function paired up with one interaction line of the screened interaction $W(\mathbf{q}, E)$. In Fig. 3.7, the classes of diagrams included in the approximation are shown. The closed loops consist of two Green's functions and it is the diagram for the polarization [6]. Following the structure of the diagrams in Fig. 3.7, the screened interaction can be written in terms of its own Dyson equation. Schematically we write it like $W = v + vPW$, with v as the bare interaction and P as the polarization. Solving for W we find $W = (1 - vP)^{-1}v$, where we identify $\epsilon = 1 - vP$ in Eq. (3.22). In this section we will approximate the polarization and use the result to extract a dielectric function.

Let us evaluate the polarization bubble, following the thesis of Nelander [91], using the non-interacting forms of the Green's functions [85]

$$G^{0,\mathrm{r/a}}_{\alpha\beta}(\mathbf{k}, E) = \delta_{\alpha\beta}\frac{1}{E - E_{\mathbf{k},\alpha} \pm \mathrm{i}\eta} \tag{3.23}$$

$$G^{0,<}_{\alpha\beta}(\mathbf{k}, E) = \mathrm{i}f(E)A_{\alpha\beta}(\mathbf{k}, E) = 2\pi\mathrm{i}f(E)\delta_{\alpha\beta}\delta(E - E_{\alpha,\mathbf{k}}), \tag{3.24}$$

where the limit $\eta \to 0$ is implied. We have used the distribution function $f(E)$, which is equal to the Fermi distribution in equilibrium, and the spectral function $A(\mathbf{k}, E)$, describing the density of states as a function of both $\mathbf{k}$ and E. The spectral function is linked to the retarded and advanced Green's functions as $A(\mathbf{k}, E) = -2\Im\{G^{\mathrm{r}}(\mathbf{k}, E)\}$. Using the Feynman rules [86] for the polarization bubble of Fig. 3.7 we apply momentum and energy conservation, add a factor $-\mathrm{i}$ for a closed fermion loop, sum over the internal degrees of freedom, and sum over spins as

those diagrams are equivalent, to find

$$P_{\alpha\beta\gamma\delta}(\mathbf{q}, E) = -2\mathrm{i}\sum_{\mathbf{k}} \int \frac{\mathrm{d}E'}{2\pi} G_{\alpha\gamma}(\mathbf{k}+\mathbf{q}, E'+E) G_{\delta\beta}(\mathbf{k}, E') \tag{3.25}$$

where the Green's functions are propagating in opposite time directions to ensure momentum conservation in the diagram. This can now be transformed from the complex contour to the real time axis using the Langreth rules [83], according to $(AB)^{\mathrm{r}} \rightarrow A^{\mathrm{r}}B^{<} + A^{<}B^{\mathrm{a}}$. The delta functions in the lesser parts of Eq.3.24 simplifies the expression, and we have for the retarded part of the polarization

$$P^{\mathrm{r}}_{\alpha\beta\gamma\delta}(\mathbf{q}, E) = 2\sum_{\mathbf{k}} G^{\mathrm{r}}_{\alpha\gamma}(\mathbf{k}+\mathbf{q}, E_\beta + E) f_{\beta,\mathbf{k}} + f_{\alpha,\mathbf{k}+\mathbf{q}} G^{\mathrm{a}}_{\delta\beta}(\mathbf{k}, E_\alpha - E) \tag{3.26}$$

where the δ's are omitted as they are already included in $G^{0,\mathrm{r/a}}$. Inserting the simple forms of the retarded and advanced propagators we reach

$$P^{\mathrm{r}}_{\alpha\beta\gamma\delta}(\mathbf{q}, E) = 2\delta_{\alpha\gamma}\delta_{\beta\delta} \sum_{\mathbf{k}} \frac{f_{\beta,\mathbf{k}} - f_{\alpha,\mathbf{k}+\mathbf{q}}}{E_{\alpha,\mathbf{k}+\mathbf{q}} - E_{\beta,\mathbf{k}} - E - 2\mathrm{i}\eta} \tag{3.27}$$

known as the Lindhard formula for the dielectric function, as we can write $\epsilon(\mathbf{q}, E) = 1 - V_q P^{\mathrm{r}}(\mathbf{q}, E)$. It is also known as the random phase approximation (RPA)[1].

Single plasmon pole approximation

To be able to describe dynamic screening at a reasonable computational cost, we use the so called SPPA to approximate Eq. (3.27) [6]. This way we can describe both W^{r} and the lesser part $W^{<}$, which is more complicated to calculate. The SPPA is best defined in terms of its inverse dielectric function

$$\frac{1}{\epsilon(\mathbf{q}, E)} \approx 1 + \frac{\omega^2_{\mathrm{pl}}(q)}{(\omega + i\delta)^2 - \omega_q^2} \tag{3.28}$$

where $\omega_{\mathrm{pl}}(q)$ is the plasma frequency which has a q-dependence in 2D [6], and ω_q is the *effective plasmon* frequency. This provides a simple expression for the dynamic

[1] This approximation keeps terms $\langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \rangle$ and $\langle a^\dagger_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}} \rangle$ in the solution of the equation of motion for the charge density, and equates them to $f_{\mathbf{k}}$ since other terms will oscillate faster and give negligible contributions [6]

screening. In this model we screen all Coulomb matrix elements $V_{\alpha\beta\gamma\delta}(q)$ by one effective band, containing the total density of electrons. The physical argument behind Eq. (3.28) is that in the long-wavelength of $q \to 0$, the dielectric function using the Lindhard formula, reduces to the classical result

$$\epsilon(\mathbf{q} \to 0, E) = 1 - V_{q\to 0} P^{\mathrm{r}}(\mathbf{q} \to 0, E) \tag{3.29}$$

$$\approx 1 - \frac{\omega_{\mathrm{pl}}^2}{\omega^2} \tag{3.30}$$

which is identical to Eq. (3.28) if $\omega_q = \omega_{\mathrm{pl}}$. The trick is then to replace the plasma frequency in the denominator with an effective pole, that should capture the physics also at $q \neq 0$. To ensure this, the dispersion of this pole frequency can be forced to respect certain sum rules. Following Ref. [6], we give the dispersions for 2D and 3D as

$$\text{2D:} \qquad \omega_q^2 = \omega_{\mathrm{pl}}^2(q)(1 + \frac{q}{\kappa}) + \nu_q^2 \tag{3.31}$$

$$\text{3D:} \qquad \omega_q^2 = \omega_{\mathrm{pl}}^2(1 + \frac{q^2}{\kappa^2}) + \nu_q^2 \tag{3.32}$$

where κ is the inverse screening length and $\nu_q = c \cdot E_q/\hbar$ is a parameter intended to take into account the pair continuum discussed below, with c being a numerical parameter. The behavior at low q was already built in the formalism, but these dispersions make sure that the low energy limit is also well behaved. For $\omega \to 0$ we have to leading order in $\mathbf{q}$ and in 3D,

$$\frac{1}{\epsilon(\mathbf{q}, E \to 0)} \approx 1 + \frac{\omega_{\mathrm{pl}}^2}{-\omega_{\mathrm{pl}}^2\left(1 + \frac{q^2}{\kappa^2}\right)} = \frac{q^2}{q^2 + \kappa^2} \tag{3.33}$$

which is exactly the static screening result in Eq. (3.22). The static screening also holds in 2D with another end result. For the long wavelength limit we see that the plasma frequency vanishes for the 2D expressions. As the quasi-2D system of the QCL definitely will go towards a 3D system in the long wavelength limit, this might pose a problem in our application. By this argument, the 3D dispersion would be a better choice. Indeed, in Paper IV this gives a better result. The dispersions for 2D and 3D are plotted in Fig. 3.8, both with (solid) and without the parameter ν_q (dashed colored lines) for the contributions from the pair continuum.

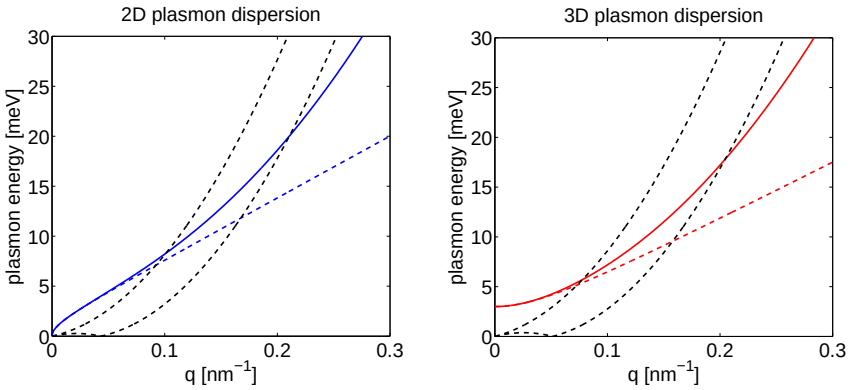

Figure 3.8: The plasmon dispersions with a numerical parameter $c = 0.3$ in Eqs. (3.31) and (3.32) are shown (solid lines), as well as with the contributions from the pair continuum turned off (dashed colored lines). Parameters for the dispersions were taken from simulations of the laser of Ref. [92] in Paper IV. The dashed black lines indicate the region of the pair continuum where the contribution from the plasmon quasi-particle is assumed to be washed out be the pair-continuum excitations. These lines are sharp at $T = 0$ but gets blurred at higher temperatures.

Pair continuum

In this section we will discuss the pair continuum and its role for the plasmon dispersions of Eqs (3.31) and (3.32). The results are summarized in Fig. 3.8 where the range of the continuum is shown, and how it influences the model dispersion. We are interested in the poles of the dielectric function in Eq. (3.27), as these will describe the natural oscillations of the electron gas. The *eigenmodes* at these poles are the collective oscillations that we refer to as plasmons. For one band only, $\alpha = \beta$, we analyze the pole structure of Eq. (3.27), and rewrite

$$P^{\mathrm{ret}}_{\alpha\alpha}(\mathbf{q}, E) = 2\sum_{\mathbf{k}} f_{\alpha,\mathbf{k}} \left[\frac{1}{E_{\alpha,\mathbf{k}+\mathbf{q}} - E_{\alpha,\mathbf{k}} - E - \mathrm{i}\eta} - \frac{1}{E_{\alpha,\mathbf{k}} - E_{\alpha,\mathbf{k}-\mathbf{q}} - E - \mathrm{i}\eta} \right] \tag{3.34}$$

in order to identify the two resonances at

$$\begin{aligned} E_{\mathbf{k}} &= E_{\mathbf{k}+\mathbf{q}} - E_{\mathbf{k}} = \frac{\hbar^2 kq}{m^*}\cos\theta + \frac{\hbar^2 q^2}{2m^*} \\ E_{\mathbf{k}} &= E_{\mathbf{k}} - E_{\mathbf{k}-\mathbf{q}} = \frac{\hbar^2 kq}{m^*}\cos\theta - \frac{\hbar^2 q^2}{2m^*}. \end{aligned} \tag{3.35}$$

For a thermalized electron gas, the electrons are limited in momenta to $\pm k_F$, where the Fermi momenta is $k_F = (2\pi n_{2D})^{1/2}$ in 2D, and $k_F = (3\pi^2 n_{3D})^{1/3}$ in 3D. Allowing only values of $k = \pm k_F$ in Eqs. (3.35) we get the upper and lower limit of the possible pair excitations, and this is what is plotted in Fig. 3.8 as dashed black lines. When the plasmon enters this regime it will be very sensitive to the large number of scattering transitions that can dephase the collective motion of the electrons. This is taken into account by the term $\nu_q = c \cdot E_q/\hbar$, and in Paper IV we use $c = 1$. One effect of this is that the pole energy quickly becomes large for increasing q, which effectively favors low energy transitions over high energy ones, as a large pole energy $\hbar\omega_q$ decreases the coupling.

Implementation

As reported in Paper IV, a GW self energy with the screened interaction treated in the SPPA was implemented in the code to study the effects of e-e scattering in a four-well THz QCL. The same self-energies was then later used in Paper VI to address the potential impact of e-e scattering in all of the structures, of various designs. In addition to the dynamic part of the self energy, there is also a static part coming from the first term in Eq. (3.28). This part will gives rise to an exchange shift, that only enter the retarded self-energies to renormalize the one-particle energy levels. This was also implemented in Paper IV. Together, these contributions to the computational model makes up a substantial part of this thesis.

The resulting GW self-energy resembles the one of for optical phonon scattering in Eq. (3.18). Energy is exchanged with a bosonic distribution providing inelastic scattering. There are three main differences: First, the scattering matrix elements are different and we show X^{SPPA} here for completeness,

$$X^{\mathrm{SPPA}}_{\alpha\alpha'\beta\beta'}(\theta) = \frac{1}{2\pi}\frac{me^2}{2\epsilon_0\epsilon_r\hbar^2}\frac{V_{\alpha\alpha'\beta\beta'}(\theta)}{q(\theta)}\frac{E_{\mathrm{pl}}(q(\theta))}{2E_{\mathrm{eff}}(q(\theta))}. \tag{3.36}$$

Secondly, the plasmons, unlike the phonons, have a dispersion, meaning that $E_{\mathrm{eff}}(q(\theta))$ entering the Green's function energy argument in Eq. (3.18) vary with wavevector q unlike E_{LO}. This makes it slightly more complicated to evaluate the self energies, which has to be done at each iteration to ensure the correct dispersion relations. Thirdly, the occupation factors $n_{\mathrm{LO}} \rightarrow n_{\mathrm{eff}}(q)$ appearing in Eq. (3.18) will depend on an effective electron temperature. This is chosen to balance the electric power dissipated into the structure, as discussed in the Appendix

of Paper VI, and thus also needs to be constantly updated. The Coulomb matrix elements, $V_{\alpha\alpha'\beta\beta'}(\theta)$ in Eq. (3.36), were calculated using the simplified scheme of Ref. [64] where only the first Fourier components of the envelope functions are used to approximate the integrals. This is very useful as it drastically reduces the computational time.

3.6 Gain saturation

In Paper I and in Ref. [93] we introduced a method to calculate both the current and intensity in QCLs under operation. As discussed below in Sec. 5.3 of Chap. 5, the gain decreases, to first order in the intensity, as

$$G(\omega) \propto G_0(\omega) - \alpha(\omega) F_{\rm ac}^2 \tag{3.37}$$

where the square of $F_{\rm ac}$ is proportional to the intensity, and $G_0(\omega)$ is the unsaturated, i.e. linear response gain. Increasing the field strength reduces the inversion, and thus the gain. In our model we are also able to resolve high intensity phenomena as long as we make the correct expansion of the system in $h_{\rm max}$ as discussed in Sec. 3.4.

To calculate the current and ac field strength under operation we simulate the gain spectrum in order to find the gain peak with respect to frequency at each bias point. This is shown for a bias of 50 mV per period for the structure of Ref. [94] in the inset of Fig. 3.10, where we find the peak gain close to 14 meV, which is in line with the experimental observations. After this, we increase the ac field strength until the gain reaches the threshold value. In Fig. 3.10 this is shown to be at an ac

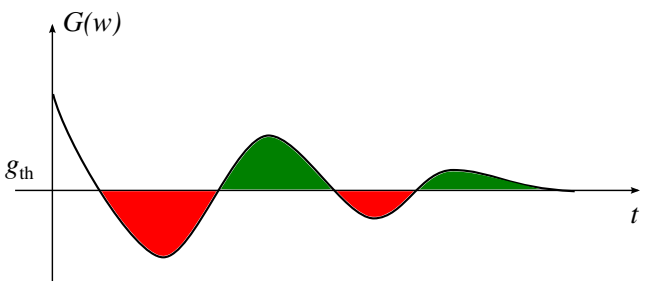

Figure 3.9: Shematic view of gain stabilizing to the threshold gain value $g_{\rm th}$. In the green regions the ac field is enhanced, which saturates the gain. When the gain is below the threshold value, red regions, we have effectively absorption, and the intensity is decreasing allowing gain to recover.

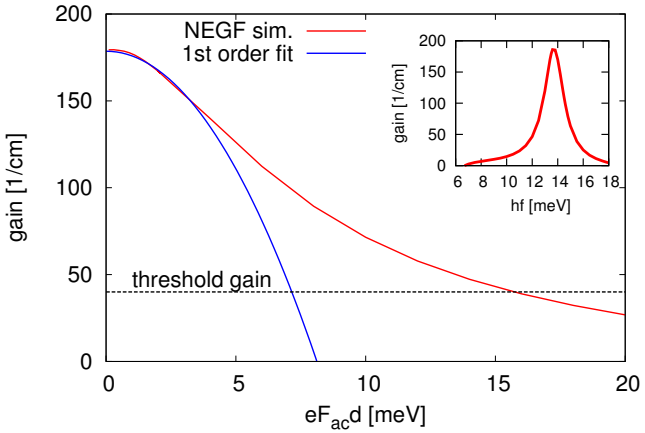

Figure 3.10: Gain calculations for the structure of Li *et al.* [94] with the NEGF model compared to a simple $\chi^{(3)}$-model with saturation linear in intensity. Inset shows unsaturated gain as a function of photon energy. Threshold gain was estimated to be 40/cm in this structure and the simulations were carried out at a lattice temperature of 77K.

field strength of $eF_{\mathrm{ac}}d = 16$ meV where we can also read off the current under operation. For comparison we show also gain according to the model of Eq. (3.37), where the gain always decreases linearly with the intensity. This is the blue line of Fig. 3.10, and it can be seen only to hold for low ac field strengths.

3.7 Summary

Using the scattering self-energies in the self-consistent Born approximation, we update the Green's function via the Dyson equation (3.8) and the Keldysh relation. The Green's functions are used to update the self-energies, until convergence is reached. The lesser Green's function is then projected down to the density matrix, from which we extract the observables that we are interested in, as discussed in Paper II. This way we calculate the time dependent current expressed as a Fourier series around the ac field frequency ω_0. From this we calculate the conductivity, and by doing this for different field strengths, we can get the gain as a function of ac field strength. Other interesting observables used in this thesis are the current components at $2\omega_0$, which we relate to second harmonic generation in Paper III.

4 Quantitative results from non-linear response simulations

4.1 Temperature dependent gain saturation

Increased lattice temperature affects the operation of especially THz QCLs in a detrimental way. Following Ref. [37] we can list the most important temperature degrading mechanisms as (i) thermally activated phonon scattering between upper and lower laser state, (ii) thermal backfilling and (iii) leakage into the continuum. In addition, there is also (iv) linewidth broadening as the impurities in the system are less efficiently screened [95]. In (i), thermally excited electron reach an energy $E^*_{\mathrm{uls}} - E_{\mathrm{lls}} \sim E_{LO}$ from where they quickly relax to the LLS. Effect (ii) consists of thermal excitations to the LLS from below, and effect (iii) is thermal excitations to the continuum of states above the barriers, possibly enhanced by the non-equilibrium LO phonon distribution built up under operation [96, 97].

In Paper I, we studied the temperature dependent gain saturation of the structure of Ref. [98] and compared to experimental data. The NEGF model, applied here without e-e scattering, includes all degrading mechanisms listed above except (iii), although we evaluate the occupation of the LO phonon mode in thermal equilibrium using the lattice temperature. The quantitative agreement is good, as shown in Fig. 4.1: At the laser frequency of 2.2 THz (9.1 meV) we simulate gain above the experimental level of losses of 18/cm for temperatures up to 120 K. Maximum operation temperature was not reported in the study of Ref. [98], however a similar structure [99] lased in pulsed mode up to a heatsink temperature of 72 K. Other similar structures grown at TU Wien lased up to 145 K in pulsed mode [100], which

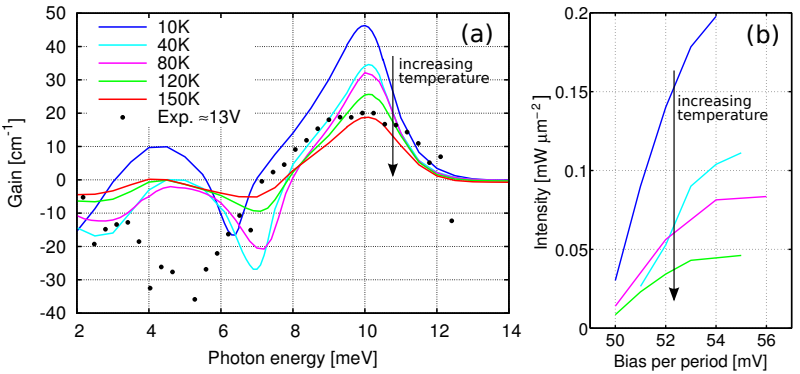

Figure 4.1: From Paper I. Collected results from simulations on the structure in Ref. [98] for a set of different temperatures. A bias of 54 mV per period was applied matching the designed value. The experimental data taken at a heatsink temperature of $\sim$ 30 K is shown as black dots.

shows that our result is in the correct range with respect to these measurements. Fig. 4.1 (b) shows how the laser intensity goes down with increasing temperature. Here we acquire intensity data by saturating the gain down to the level of the losses by increasing the ac field strength, as outlined in Sec. 3.6.

Analyzing again the simulation data that Paper I was based upon, it is possible to assess the most important temperature degrading mechanisms, referring again to the list in the first paragraph. The gain is found to decrease quicker than the inversion, implying that linewidth broadening (iv) is important. Comparing to Eq. (2.22), which is $G(\omega) \propto \Delta n_{12}|z_{12}|^2/\Gamma_{12}$ at resonance, this underestimates the gain, if we approximate the linewidth with the state lifetimes. Inspecting Fig. 4.1, it is however possible to observe that the linewidth is not increasing strongly with temperature. This is due to the correlated scattering environments in this vertical design [101], which explains the slower decrease of the peak gain. The inversion is not very sensitive to thermal backfilling (ii) at these temperatures, it is rather an increased scattering from upper to lower laser state that is responsible for the decrease in inversion, partly via thermally activated phonon scattering (i).

4.2 Quantitative results under operation

To benchmark our model, and to asses its reliability in the THz range, we performed a large study of 16 devices reported in the literature during the last ten

years. Within this group we managed to include four distinctly different active region design classes, that all utilize the ultra-fast optical phonon scattering to ensure inversion. More importantly, samples from five different MBE labs were considered. The lasers studied consist of samples from previous studies performed in our group, as well as new samples, included in order to make the study span over more different design classes and growth places.

The method that was applied in this study was exactly the same for all samples and it employed the most recent model improvements. One example being the new parametrization of E_{typ} as discussed in Sec. 3.3. This and all other simulations parameters, such as the interface roughness (IFR) distribution, were left unchanged. Each sample was simulated using the nominal sheet doping density and layer sequence as given in the respective journal publication. The threshold gain g_{th} was treated differently for metal-metal waveguides and the semi-insulating surface plasmon (SI-SP) waveguides. For the metal-metal waveguides one can approximate the reflectivity by the impedence mismatch method [102]. The wave impedence goes like the electric over the magnetic field. In MM waveguides, the light is polarized with its electric field in the growth direction z, and with the magnetic field of the TM mode [4] in the y direction, along the width of the waveguide. If the wave is confined in any of these directions, the impedence will differ from its constant value $Z(\lambda) = \sqrt{\mu/\epsilon}$. Using this poor man's argument, the wave impedence will

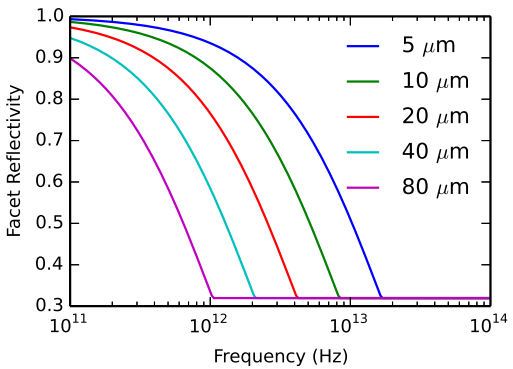

Figure 4.2: Reflectivities for different waveguide heights as a function of frequency, using the simple model of Eq. (4.1). An infinite width was assumed. For large heights and frequencies all approach the Fresnel losses of 0.32 for an GaAs/air interface.

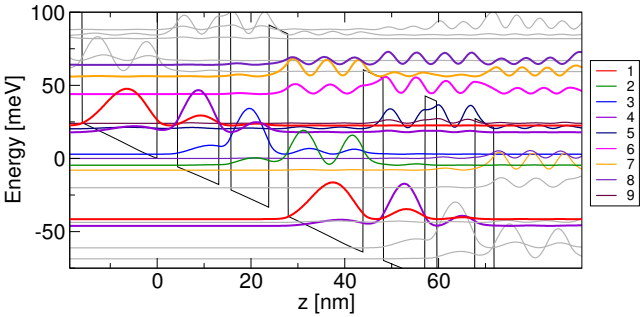

Figure 4.3: Shown here are the Wannier-Stark states for the structure of Ref. [29] for a bias of 64 mV per period at 77K lattice temperature. This is slightly after optimal injection. States 6-8 are emphasized as their inclusion are important for current continuity as shown in Fig. 4.4.

approximately follow

$$Z(\lambda) = \frac{1}{n_r}\frac{\min(\lambda, h)}{\min(\lambda, w)} \cdot Z_0 \tag{4.1}$$

where h and w the waveguide height and width, respectively, and $Z_0 = \sqrt{\mu_0/\epsilon_0}$ is the vacuum impedence. From this we evaluate the reflectivity $R = |Z - Z_0|^2/|Z + Z_0|^2$ [102], plotted in Fig. 4.2 for different heights h of infinitely wide waveguides. Compared to Ref. [103], where this was solved using the standing wave ratio using the exact eigenmodes of a standard cavity, we find qualitatively the same results.

In the NEGF model, there are a number of accuracy parameters that we have to converge the results against. The most important are the number of states in each period N_ν and the number of adjacent periods to the central one N_{per}. To get a measure of the accuracy we normally look at the continuity of the current. As the practical inclusion of a complete basis is impossible there will be some fluctuations over the period [41], which we average over in order to get the current density. In Fig. 4.3 we show the simulated Wannier-Stark states at a bias of 64 mV per period for the structure of Ref. [29]. The structure is biased slightly beyond optimal alignment of injector (state 1) and upper laser state (state 4). For this situation we plot the current as a function of growth direction z in Fig. 4.4, where we have converged the simulations with different parameters. The lower part clearly shows the need to include at least two adjacent periods in the simulations. In the upper part, we see that seven states is a minimum, while increasing to nine gives a slight increase in current. Based upon such a continuity study, it can be concluded that $N_{\text{per}} = 2$ and $N_\nu = 7$ constitute a minimum state space criteria to capture the

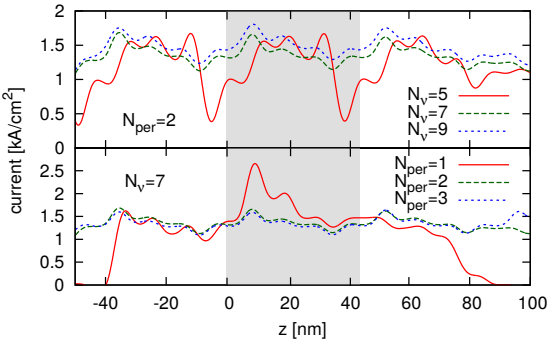

Figure 4.4: Current continuity for varying number of states per period N_ν and adjacent periods N_{per} included at the for the situation in Fig. 4.3. The upper section has a constant $N_{\mathrm{per}} = 2$ while the lower has fixed $N_\nu = 7$. The shaded area indicates the central period.

essential physics. From Fig. 4.3 it can be seen that the states 6-8 form a band of states, and the results of Fig. 4.4 indicate that they are important to correctly describe the coherent tunneling over the injection barrier, between state 1 and 4. High lying states can often originate from a period two steps downfield, why the inclusion of more neighboring periods becomes important as well.

In all papers of this thesis we use a course grain resolution in both E and $\mathbf{k}$. For the samples in Paper VI the resolution was chosen to ~ 0.3 meV based on previous experience, for both ΔE and ΔE_k. In short, the study of the large number of samples

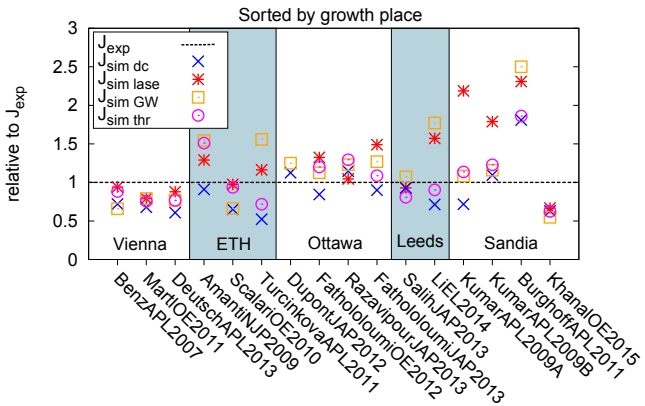

Figure 4.5: From Paper VI. A graphical display of the collected results of Paper VI. Here, the samples are referenced by a shorthand notation. For full references, please see the paper. We sort the samples according to growth place, which provides a possible explanation of the differences in simulated and experimental current.

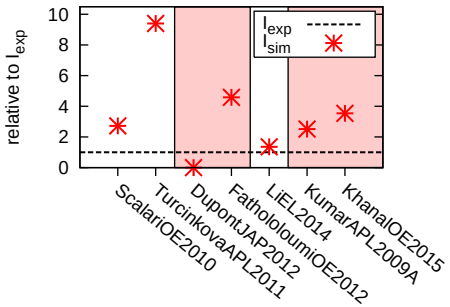

Figure 4.6: Simulated intensities relative to the experimentally measured. Here the samples with more than a tenfold disagreement have been left out.

in Paper VI, with results shown in Fig. 4.5, shows that our model is reliable within each group. If the structure for example is grown in Vienna, we expect to have good agreement although we normally underestimate the measured current. For samples grown at the National Research Council of Ottawa, we instead slightly overestimate the current, although overall agreement is good. As the differences with respect to lab (as well as the difficulties in reproducing THz QCL designs overall) are well known, the fact that we are consistent for samples from each lab is indeed a good demonstration for the validity of the model. Here it should also be noted that we show all our results and not only the structures with good agreement. As an additional point we were also able to show that different IFR distributions have only small effect on the current under operation. This is a valuable contribution to the understanding of how this scattering mechanism influences the operation of THz QCLs and can be added to the earlier results of our group on this topic [104, 105]. While the currents did agree well for most samples, the intensities were largely overestimated. In Fig. 4.6 we show the samples for which we achieved comparable numbers (in many publications the intensity was never measured). Here we convert the experimental power into waveguide intensity for the comparison to simulated values. This is done under the assumption of a 30% collection efficiency, which is a standard estimate for THz QCLs [106, 107, 108]. The reason for the low collection efficiency of the MM waveguides is the poor quality of the beam pattern due to the subwavelength facet dimensions [30]. There are a few reasons why the measured intensity would be overestimated by our simulations. First, we assume that the total area of the waveguide facet would be active and supply the peak intensity. Secondly, higher order lateral modes with a lower output coupling could be ignited in the waveguide [109], saturating the gain but

with little increase in power output. Thirdly, the facet transmittivities could be worse than the ideal values shown in Fig. 4.2 and we might also underestimate the threshold gain since this is seldom measured. Fourthly, we neglect the back-reflected wave from the facets, discussed in Paper III, which could be large as the reflectivity is over 70% for low frequencies (see Fig. 4.2).

4.3 Simulating second harmonic generation

In centro-asymmetric systems, the heterostructure of quantum wells can exhibit a strong second-order susceptibility $\chi^{(2)}$ [111] which for example is used in difference frequency generation (DFG) to down-convert mid-IR radiation to THz [21]. The same physical principle is responsible for SHG, also known as *frequency doubling*. It was shown already in 2003 [112] that the QCL is capable of acting both as the pump and the nonlinear optical converter for SHG. This is the same idea that was recently used to realize DFG devices that operate up to room temperature [23], where two active regions emitting at two slightly different energies are grown in the same active region, and the nonlinear element is built into the design. A similar concept with a ladder structure was used in an optimized mid-IR structure aimed

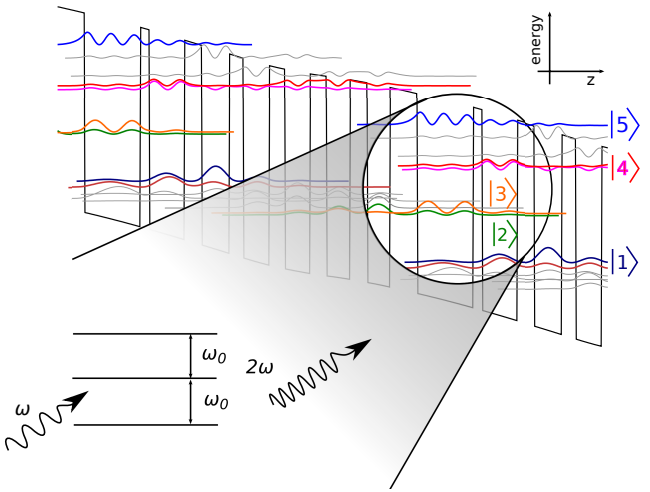

Figure 4.7: Illustration of the second harmonic generation (SHG) process in the optimized structure of Ref. [110]. The laser transition from state 3 to 1 with frequency ω_0 acts as a pump and the ladder structure of state 4 and 5 are designed to improve the conversion efficiency.

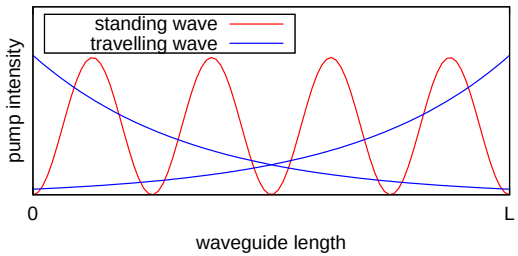

Figure 4.8: Pump intensity sketched for the two extreme cases of total reflection (red) and near unity transmission (blue). L is the laser cavity length.

at increasing the conversion efficiency for SHG [110]. The structure is shown in Fig. 4.7 where the pumping is done by the inverted population difference between state 3 and state 1. The ladders 1-3-4 and 3-4-5 will provide a large second-order susceptibility for SHG. In Paper III we present results were we model this process with good accuracy.

From the NEGF formalism we expand our time dependent observables in a Fourier series with the pumping frequency as the fundamental frequency. We use the current oscillating with the fundamental frequency to calculate the gain coefficient and add higher orders to the expansion until we reach convergence. The higher order terms carry information about the polarization at the higher harmonics, such as the second and third, in terms of the fundamental frequency. We can thus extract also the current term $J(2\omega_0)$, and relate this to the field generated in the cavity at this frequency.

As a simple model to estimate the intensity generated by this current inside the waveguide, we view the second order current as a series of antennas. The mathematical detailed are outlined in the Appendix of Paper III. One important aspect of this, is that the intensity of the SHG wave is proportional to the square of the pump intensity, and it is thus important to know the spatial form of the pump wave. Here we can consider two extreme cases with respect to the facet reflectivity as sketched in Fig. 4.8; if we have perfect reflection, we would have a standing wave modulating the second harmonic intensity, and for near total transmission, we would have a traveling wave that would start over every time it hits the facet. The two approaches differ by a factor of two for pump intensity (ω_0) for the pump power and a factor 4 for the SHG power ($2\omega_0$), with higher powers for a traveling wave as the back-reflected wave increases gain saturation. Our results are in good

agreement and slightly overestimates the experimental results. In our approach we do not calculate the full modal overlap of the fundamental and second order modes for the waveguide geometry [110], which could explain the overestimation of the SHG intensity.

5 Revealing physics in quantum cascade lasers

The full quantum approach described in Chap. 3 provides us with the Green's function and self-energies, but these abstract objects on their own cannot teach us much about the physics of the QCL. It is instead the derived quantities like the density matrix that can give us substantial feedback and knowledge, that we can connect to measurable observables. In order to understand the results of the full model, it is often helpful to compare to simpler approaches to see if that can explain the result.

In this chapter we give a number of examples where the full quantum results have been compared to more straightforward approaches, and what we have been able to learn in those cases. We will also discuss the possibility of parametrizing the response of the QCL with analytic functions, an advantageous method using physically based interpolation schemes in order to span a wider parameter space.

5.1 Second harmonic generation from perturbation theory

In the previous chapter we discussed how the current at the second harmonic frequency could be calculated within the NEGF formalism. The average current over one active region was used to estimate the generated power in the waveguide. Another way of approaching the problem, which also might give some insight, is to use the results from time-dependent perturbation theory (TDPT) from Chap. 2. The polarization is then retrievable as a function of field strength if we are pro-

vided the dipole matrix elements and the occupations of the respective states. The current is the time derivative of the polarization, which we express as a Fourier series

$$P(t) = \epsilon_0 \epsilon_r \sum_{\omega_i} \chi(2\omega_i) E^2(\omega_i) \mathrm{e}^{-\mathrm{i}2\omega_i t}, \tag{5.1}$$

as discussed in relation to the current in the previous chapter. Here the sum should include also the negative frequencies in order to correctly build the cosine field. The time derivative reads

$$J(t) = \dot{P}(t) = -\omega_0 \epsilon_0 \epsilon_r F_{ac}^2 \Big[\Re\{\chi^{(2)}(2\omega_0)\} \sin(2\omega_0 t) \tag{5.2}$$

$$- \Im\{\chi^{(2)}(2\omega_0)\} \cos(2\omega_0 t) \Big], \tag{5.3}$$

where we have, as always, assumed a classical light field $E(t) = F_{\mathrm{ac}} \cos(\omega_0 t)$ which gives $E(\pm\omega_0) = F_{\mathrm{ac}}/2$. The second order susceptibility was derived in Chap. 2, and it can now be used here, with the results of the Green's function calculations as input. If we plug in the occupations and the dipole matrix elements that we get if we diagonalize the superlattice Hamiltonian, H_{SL} in Eq. (3.1), at the bias points we are interested in, we can calculate the real and imaginary part of the susceptibility. This yields an approximate result that we can compare to the full simulation in Fig. 5.1. Here we plot *particle current*, why the sign is reverse compared to what we normally are used to.

From this result we can say a few things about our approximation, which is quite drastic when we compare to the results of the NEGF calculations. There, the current at the second harmonic is calculated directly from the elements of the density matrix, including the coherences (the off-diagonal elements). In our approximation we have only included the steady state occupations, and thus assumed that the density matrix is diagonal in our choice of basis. As we will discuss in the next section, this is not necessary the case. We have also approximated the transition linewidths to $\gamma_{ij} = (\Gamma_i + \Gamma_j)/4$, which is less than the mean of the state lifetimes Γ_i, in order to take correlation effects into account [101].

Albeit the harsh approximation, the results still show some similarities. The phase of the oscillations is almost kept, especially for the cosine signal, and the amplitudes are not that far off although partly shifted. We conclude that the qualitative result is within reach with the second order susceptibility calculation, but that the quantitative predictive power demands a full quantum treatment of the coherences and the transition linewidths.

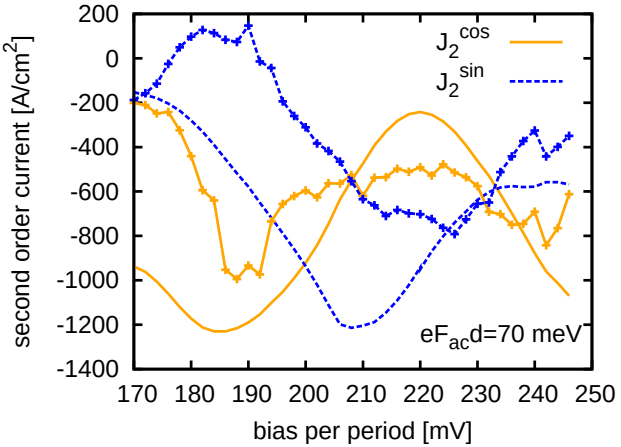

Figure 5.1: From Paper III. Second order susceptibility calculations compared to results from NEGF for the device reported in Ref.[110] at an intensity of the pump field of $eF_{ac}d = 70$ meV. Here the sine and cosine components are separated out. Dotted lines correspond to the approximate calculations.

5.2 Gain characterization

The linear response gain expression, Eq. (2.22) derived in Chap. 2, contains three important parameters; the inversion Δn, the linewidth Γ and the dipole matrix element z_{ul} between the ULS and the LLS. In order to maximize gain we need to have both high inversion and a large dipole matrix element, while we keep a narrow linewidth. If scattering is strong, the lifetimes of the states will decrease, broadening them in energy. This will in turn increase the linewidth Γ of the gain transition, although there is a reduction if the scattering environments of the ULS and LLS are correlated [101]. The scattering might also decrease the tunneling, which might reduce injection and limit inversion Δn. One example of this was shown in Paper VI where limited injection at higher scattering strengths are believed to additionally limit inversion.

The scattering environment usually does not affect the dipole matrix elements z_{ul}, as they in general are calculated using energy eigenstates. After converging the NEGF simulations, the approximate energy eigenstates can be calculated by diagonalizing the superlattice Hamiltonian $\hat{H}_{\mathrm{SL}}$ in Eq. (3.1), including also the static part of the meanfield term $\hat{H}_{\mathrm{MF}}$. Denoted Wannier-Stark (WS) states, they are approximate as they do not contain the perturbative part $\hat{H}_{\mathrm{scatt}}$, although we

incorporate some of the scattering renormalization by adjusting the energy eigenvalues by the real part of the retarded self-energy [105]. Neglecting $\hat{H}_{\text{scatt}}$, the WS states are not very sensitive to increased scattering. We can find an alternative description of the system if we instead of the WS states look at the eigenstates of the lesser Green's function [113]. We remind ourselves that this object is related to the density matrix as

$$\rho_{\alpha\beta}(\mathbf{k}) = -\int \frac{\mathrm{d}E}{2\pi \mathrm{i}} G^{<}_{\beta\alpha}(\mathbf{k}, E). \tag{5.4}$$

The density matrix is more sensitive to scattering than the energy eigenstates; if dephasing times decrease, the coherences carrying current across the barriers will also decay, which indeed will be reflected in the density matrix. It is however periodic in the module index, and thus we cannot distinguish the ground state in one period from the one in the next. It would thus give us periodic, Bloch-like functions if we were to diagonalize it, while we would rather have a localized basis. If we diagonalize the lesser Green's function instead, we will get a set of states that are both localized and dependent on the scattering environment. In addition, the eigenvalues correspond to occupations, so by diagonalizing all matrices at $\mathbf{k} = 0$ for example, we get energy-resolved information and an accurate measure to where in energy we have high occupation probabilities. This is more precise than the eigenenergies of the WS states, where changes to the spectral function and the actual non-equilibrium distribution function do not enter. An example is shown in Fig. 5.2 (a) where we show the eigenvalues from diagonalizing $G^{<}(\mathbf{k} = 0, E)$ of the superlattice of Paper V. This superlattice is proposed to exhibit inversion gain in a region of positive differential conductance and the principle of operation

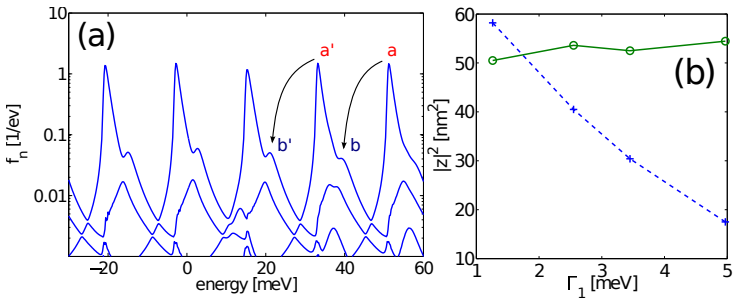

Figure 5.2: From Paper V. (a) Eigenvalues of $G^{<}(\mathbf{k} = 0, E)$ as a function of energy. The level crossings are missing as the diagonalization orders the eigenvalues in size. (b) Computed dipole matrix elements squared, using either WS states (green solid) or eigenstates to the lesser Green's function (blue dashed).

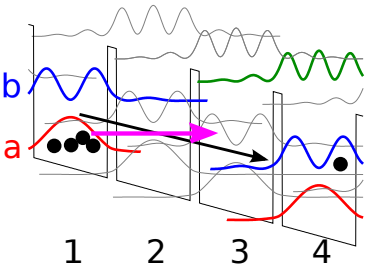

Figure 5.3: Principle of operation of the superlattice design proposed by A. Andronov [12]. Here we plot the Wannier states of the ground, excited and second excited states in red, blue and green respectively. The gain transition is indicated by a black arrow. The tunneling transition ensuring positive differential conductivity, from state a of well 1 to state b of well 3, is shown by a magenta arrow.

is shown in Fig. 5.3. In the logarithmic scale of Fig. 5.2(a) we can actually read out where the electrons are located in energy. Here we can directly observe the inversion in energy from the ground state to the second excited state two periods in the downfield direction as shown in the band diagram of the operating principle in Fig. 5.3. The eigenstates of $G^<(\mathbf{k}, E)$ are also functions of energy, but do not vary on the same quick scale as the eigenvalues. If we pick the eigenstates at the density maxima, we can compute new dipole matrix elements. These are compared to the ones calculated from WS states in Fig. 5.2 (b), where it is shown how the increased scattering proves to be devastating to the operating principle. As more impurities are introduced into the system, scattering increases and counteract the effect of more carriers and inversion. This is the reason why gain is only slightly increasing with doping in the calculations of Paper V.

At low temperatures, where screening is efficient, we could confirm the principle of operation of a device based on gain from quantum coherence. Using the lesser Green's function and the basis states, we can resolve the gain from the NEGF simulations in position and energy using the expression

$$G(z,E) = \frac{e^2}{2\pi c\epsilon_0\sqrt{\epsilon_r}d}\sqrt{\frac{E_g(z)}{2m_c(z)}}\frac{1}{eF_{\mathrm{ac}}d}$$
$$\times \sum_{\alpha\beta} \Re\left\{g^<_{\alpha\beta,1}(E) + g^<_{\alpha\beta,-1}(E)\right\}\left(\Psi^c_\alpha(z)\Psi^c_\beta(z) - \Psi^v_\alpha(z)\Psi^v_\beta(z)\right), \qquad (5.5)$$

where we have used $g^<_{\alpha\beta,b}(E) = \int \mathrm{d}E_k G^<_{\alpha\beta,b}(E_k, E)$, and c, v corresponds to conduction and valence band components of the wavefunction $\Psi(z)$, respectively. Here, the z dependent energy gap E_g as well as the conduction band effective

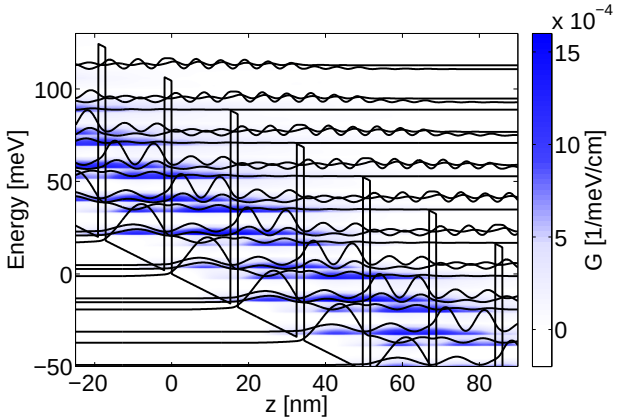

Figure 5.4: From Paper IV. Using Eq. (5.5) we can resolve the gain at $\hbar\omega_0 = 13$ meV as a function of position and energy. The plot shows that gain comes from the delocalized coherences stretching over several periods.

mass m_c enters. For details on the two-band model, please consult the thesis of Franckié [81]. At the proposed operation point at 18 mV per period we show the resolved gain in Fig. 5.4. From these results we can confirm the principle of operation, and see that the gain is dependent on long range quantum coherences in the structure.

5.3 Parametrization of the QCL response

In our current implementation discussed in Chap. 3 we use the same frequency ω_0 both to pump and probe our system. The addition of an independent probe frequency would require the introduction of a third energy in the Green's function (currently we have the E from the time difference and h for the harmonic index), which is beyond the scope of this thesis. One alternative is to map the results of the NEGF calculations onto a simpler model of the gain medium, and instead probe that model system with multiple frequencies.

A simple model also allows us to make quick comparisons with analytic expressions, to see when our results are in line with simpler methods or if a full quantum treatment is necessary. It is also motivated by the possibility to make simulations on a more macroscopic level, combining active regions of different type or situations where the active region is inhomogeneously biased. In such situations the

current needs to be conserved but the biases can be different, and an ac field might also be present. To simulate such systems, we would like to know the current and gain as a function of bias, ac field strength and frequency of the ac pump field. As the parameter range quickly gets very large, a clever interpolation scheme is useful, so that the NEGF scheme does not have to be consulted for each point. In this section we want to present analytic functions that captures the essential behavior of THz-QCLs *under operation*. In this regime we vary both the probe frequency $\hbar\omega_0$, the bias $eF_{\mathrm{dc}}d$, as well as the ac field strength $eF_{\mathrm{ac}}d$, where we have used the period length d to convert the variables into units of energy.

Two level gain medium

Considering a simplified two level system of Fig. 5.5, where we have an injecting current J_{inj} into level 2, scattering lifetimes τ_1, τ_2 for level 1 and 2, respectively, and τ_{21} for scattering from 2 to 1. A laser field closely resonant to the energy difference E_{21} will induce a transition rate according to Fermi's golden rule [44] as

$$\begin{aligned}
\gamma_{12}(\omega)F_{\mathrm{ac}}^2 &= \frac{2\pi}{\hbar}|V_{21}|^2\delta(E_{21}-\hbar\omega) \\
&= \frac{2\pi}{\hbar}\left|\frac{ez_{12}F_{\mathrm{ac}}}{2}\right|^2\delta(E_{21}-\hbar\omega) \\
&= \frac{2\pi}{\hbar}\left|\frac{ez_{12}}{2}\right|^2 F_{\mathrm{ac}}^2\frac{1}{2\pi}\frac{\Gamma}{(E_{21}-E)^2+\Gamma^2/4} \qquad (5.6)\\
&= \frac{2\pi}{\hbar}\left|\frac{ez_{12}}{2}\right|^2 F_{\mathrm{ac}}^2\mathcal{L}(E_{21},\Gamma) \qquad (5.7)
\end{aligned}$$

where we separate out F_{ac}^2 explicitly from the rate, and V_{21} is the matrix element which is calculated using an ac field of $F(t) = F_{\mathrm{ac}}\hat{\mathbf{z}}\cos(\omega_0 t)$. As in Chap. 2 we

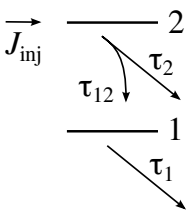

Figure 5.5: Two level system where the ULS is pumped by a current density J_{inj}. The total lifetime of the upper laser state is τ_2 whereas the time to go from 2 to 1, τ_{21}, satisfies $\tau_{21} > \tau_2$ per definition.

have replaced the δ-function with the more physically sound Lorenztian $\mathcal{L}(\Delta E, \Gamma)$. Under irradiation we can establish the rate equations, following Ref. [4],

$$\frac{\mathrm{d}n_2}{\mathrm{d}t} = \frac{J_{\mathrm{inj}}}{e} - \frac{n_2}{\tau_2} - \gamma_{12}(\omega)F_{\mathrm{ac}}^2(n_2 - n_1) \tag{5.8}$$

$$\frac{\mathrm{d}n_1}{\mathrm{d}t} = \frac{n_2}{\tau_{21}} - \frac{n_1 - n_1^{\mathrm{th}}}{\tau_1} + \gamma_{12}(\omega)F_{\mathrm{ac}}^2(n_2 - n_1) \tag{5.9}$$

where $E = \hbar\omega$ is the photon energy of the ac field, and the second term of the equation for n_1 makes sure that the population reaches its thermal equilibrium value n_1^{th} in the absence of an injecting current (and ac field). For steady state, we reach the solution

$$n_2 = \tau_2\left[\frac{J_{\mathrm{inj}}}{e} - \gamma_{12}(\omega)F_{\mathrm{ac}}^2(n_2 - n_1)\right] \tag{5.10}$$

$$n_1 = n_2^{\mathrm{th}} + \tau_1\left[\frac{\tau_2}{\tau_{21}}\frac{J_{\mathrm{inj}}}{e} + \frac{\tau_{21} - \tau_2}{\tau_{21}}\gamma_{12}(\omega)F_{\mathrm{ac}}^2(n_2 - n_1)\right] + n_2^{\mathrm{th}} \tag{5.11}$$

which yields the inversion

$$\Delta n_{21} = \tau_2\frac{\tau_{21} - \tau_1}{\tau_{21}}\frac{J_{\mathrm{inj}}}{e} - n_1^{\mathrm{th}} - \left(\tau_1 + \tau_2\frac{\tau_{21} - \tau_1}{\tau_{21}}\right)\gamma_{12}(\omega)F_{\mathrm{ac}}^2\Delta n_{21} \tag{5.12}$$

where the first two terms is what we would have for vanishing ac field strength. From these terms we can identify the criteria for inversion, namely that $\tau_{21} > \tau_1$, under the assumption that n_1^{th} is small compared to J_{inj}/e. This shows that gain is also possible if $\tau_2 < \tau_1$, as long as the scattering from 2 to 1 is small. The third term gives the saturation of the inversion, where the factor including the scattering times again carry very important physical information. As the ratio τ_1/τ_{21} should be minimized for a good QCL design, it is safe to assume that $(\tau_{21} - \tau_1)/\tau_{21} \approx 1$ which simplifies the expression. In this expression we find that the *gain recovery* time is $\tau_1 + \tau_2$, that is, the time scale on which the both gain saturates and recovers. Following Ref. [45], we can relate the linear response gain for a two level system in Eq. (2.22), to the γ-factor as

$$G_{12}(\omega) = \omega\frac{2}{c\epsilon_0\sqrt{\epsilon_r}d}\gamma_{12}(\omega)(n_2 - n_1). \tag{5.13}$$

Recasting Eq. (5.12) into $\Delta n_{21} = \Delta n_{21}^0 - (\tau_1 + \tau_2)F_{\mathrm{ac}}^2\Delta n_{21}$, the gain including saturation can be expressed, following M. Franckié in [114], as

$$G(\omega) = \frac{G^0(\omega)}{1 + \bar{\tau}\gamma(\omega_0)F_{\mathrm{ac}}^2} \tag{5.14}$$

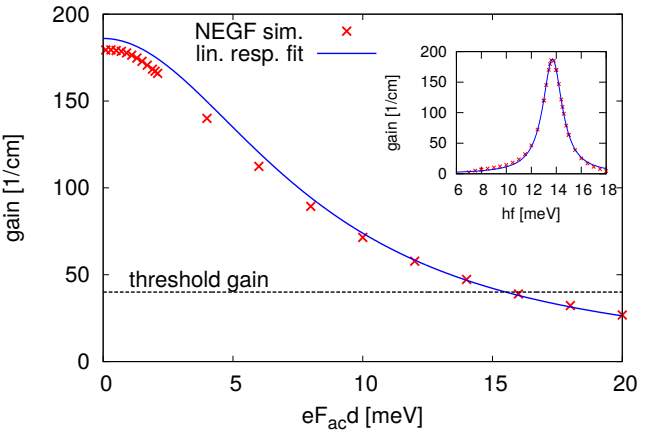

Figure 5.6: Linear response gain and saturation from NEGF simulations fitted by Eq. (5.13) (inset), and Eq. (5.14), similar to Fig. 3.9. From the fitted saturation we can extract the effective lifetime $\bar{\tau}$.

with $\bar{\tau} = \tau_1 + \tau_2$ as an effective lifetime for the transition. Here, $G^0(\omega)$ is the unsaturated linear response gain. Please note, that we have now separated the pumping ac field at frequency ω_0 from the probe at ω, which means that we can now probe the saturated gain at a chosen frequency.

If the denominator is expanded to first order in the intensity, the term with the F_{ac}^2 dependence goes like $-G^0(\omega)\bar{\tau}\gamma(\omega_0)F_{\mathrm{ac}}^2$. Both $G^0(\omega)$ and $\gamma(\omega_0)$ carries the dipole matrix element squared, which in total gives a dipole matrix element to the power of four, as discussed in Chap. 2. This shows that the gain saturation is indeed a $\chi^{(3)}$-process.

Under the assumption that the bottle neck for current is the inverted population at the laser transition, the total current will be a sum of stationary and induced current following

$$J(\omega_0) = J_0 + e(n_2 - n_1)\gamma_{12}(\omega_0)F_{\mathrm{ac}}^2 \tag{5.15}$$

where we now again have only one frequency ω_0 which saturates the system. The expression can readily be generalized to many modes of finite intensity, yielding a sum over field strengths at different frequencies ω_i, each with $\gamma_{12}(\omega_i)$. Identifying

the gain in the expression for current, using Eq. (5.14), gives

$$J(\omega_0) = J_0 + edG(\omega_0)\frac{c\epsilon_0\sqrt{\epsilon_r}}{2}\frac{F_{\mathrm{ac}}^2}{\omega_0}$$
$$J(\omega_0) = J_0 + edG(\omega_0)\frac{I(\omega_0)}{\hbar\omega_0} \tag{5.16}$$

where the last part of the first line can be rewritten as the Poynting vector giving the intensity $I(\omega_0)$ at ω_0 in the final expression. This shows nicely how a photon density $I(\omega_0)/\hbar\omega_0$ in the cavity contributes to the current with the intensity dependent factor $dG(\omega_0)$. As the gain that enters Eq. (5.16) is the saturated gain corresponding to the ac field strength F_{ac}, there will be a limit value to the current using this approximation at high intensities, found by combining Eq. (5.13) and (5.15) into

$$J(F_{ac} \to \infty) = J_0 + \frac{e\Delta n_{12}^0}{\bar{\tau}} \tag{5.17}$$

where Δn_0 corresponds to the linear response inversion. To give an example, let us revisit the structure of Ref. [94], previously presented in the context of gain saturation in Fig. 3.9 in Chap. 3. The linear response gain can be fitted to the full NEGF data with excellent agreement as displayed in the inset of Fig. 5.6. By fitting the intensity dependent gain in Fig. 5.6 to Eq. (5.14), the effective lifetime $\bar{\tau}$ of 2.5 ps is found. Using the inversion at low intensity Δn^0, Eq. (5.15) provides the limit value of 0.7 kA/cm^2 for the stimulated emission current. This fits very well with the data in Paper VI where we show additional results for this structure, and find that the current increases from 0.5 to 1.1 kA/cm^2 under lasing.

Mean-field Stark shift

The THz QCL design based on three quantum wells is of great interest as it has so far outperformed the other designs at high temperature. We have simulated both the present record device that lased up to 200 K [29] as well as the previous record holding device reaching 186 K [115]. For the latter one, when we fitted the gain spectra with Lorenztian profiles we observed a Stark shift of the lasing frequency as bias was increased. The physical reason for this shift is simple. As the bias is increased, the level spacings increase in total, and the difference between ULS and LLS gets larger. Such as shift is thus normal and seen in most lasers, however in

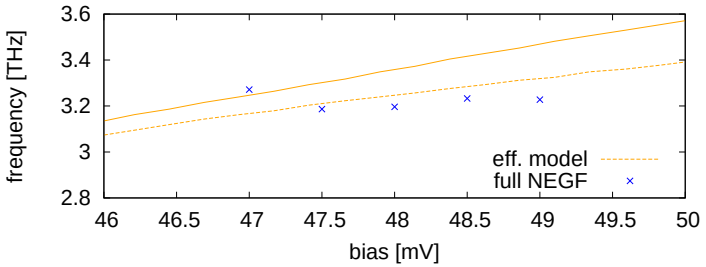

Figure 5.7: Calculations using the linear response gain compared to full NEGF simulations under laser operation of the structure of Ref. [115]. Dashed (solid) line shows the result with (without) the mean-field Stark shift.

the NEGF simulations at high intensity, this shift could not be seen. Instead the laser frequency was stable with bias as shown in Fig. 5.7. The simulations with the NEGF code were done by iteratively stepping up in bias. At each point the frequency yielding maximum gain was found, then intensity was increased until gain saturated to the level of the losses. The final intensity was then carried over to the next bias point. Using this scheme it is possible to find the actual lasing frequency at high intensity self-consistently.

Here the design of the laser active region is important to consider, and the occupations of the states at low compared to high intensity. By employing a capacitor model, we can estimate the electric field shift as carriers are transferred from the ULS to the LLS. The increasing number of carriers in the downfield region (the lower laser state) as intensity increases, will locally balance the increase in bias when addition current is pushed through the structure. We translate this into an effective frequency shift of the laser transition as

$$\hbar\omega = \hbar\omega_o + (\Delta n - \Delta n_0)\frac{ea}{2\epsilon_0\epsilon_r}$$

where a is the mean separation of carriers and a measure of the diagonality of the transition, and $\hbar\omega_0 = \hbar\omega_{ul}$ in linear response, and Δn and Δn_0 are the inversion at high and low intensity, respectively. As we model the saturation of inversion in our parametrization we can directly relate this to the frequency shift. Employing this shift for the simulations of Ref. [115], using the inter-well distance of the ULS and LLS wells for a we get a better agreement with the full calculation as shown in Fig. 5.7 by dashed lines.

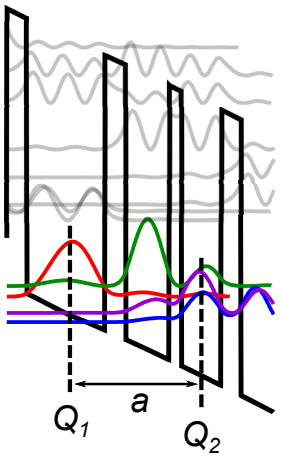

Figure 5.8: A sketch of the three-well active region of Ref. [115]. The electric field is evaluated using a capacitor model with distance a.

Conclusion

With the mean field shift, gain saturation and stimulated emission current we have a way of interpolating the NEGF simulation data in a physically sound way. This parametrization opens up possibilities of simulating active regions containing more than one type of design, to look at domain formation in lasers under operation, where the field distribution is inhomogeneous, and also pump probe situations, where the effects on the gain spectrum away from the pump frequency can be studied.

6 Summary & Outlook

In this thesis we have presented developments in several directions of modeling quantum cascade lasers under high intensities. We have shown that we can simulate the non-linear response and retrieve quantitative results for different optical effects such as second harmonic generation and gain saturation. Both these effects have been related to simpler approaches, based on the discussion in Chap. 2 which helped us to reveal the important physical effects. In the case of the superlattice studied in Paper V, this made us realize that the strongly scattering dependent gain could not be explained through the standard approaches.

The quantitative modeling has also been extended by a rudimentary form of GW where the screened interaction is included via the single plasmon pole approximation. The issue of predictivity will probably not be resolved until all relevant scattering processes, including the e-e scattering, are taken into account on the same order. In the case for our implementation of the NEGF method, this will require an extension of the theory to go beyond the rudimentary form of GW that was introduced in Paper IV. To do this one would either implement a form RPA which was outlined already in the thesis of Nelander [91], or a self-consistent GW approach where the non-equilibrium Green's functions enter the polarization bubble at each iteration.

Although the development of high-temperature THz QCL has seen very modest development during the last decade, there is still a large parameter space to probe. Tasks for a general optimization algorithm could be the optimization of certain material parameters for a given wavelength. For the material system of $Al_xGa_{1-x}As/GaAs$, the Al content can be varied freely without introducing any strain in the material, and if x is kept below $\sim$ 40%, the influence of the X-

valley of Al should be very limited [116]. This provides a optimization parameter that could be calculated for each frequency and design class. More generally, one could let the barrier height run completely free, in order to find the best material for THz QCLs, which would give the growing teams something to take aim for. This, and specialized active region optimization requires a quick way of asserting the quality of each generation of designs in the optimization algorithm. One interesting pathway is to combine the recent development of density matrix models [68, 117, 118] with the NEGF approach to get a solver that might not be perfectly predictive in all cases, but at least has all the relevant physical processes, combined with a reasonable execution time.

Acronyms

CBO conduction band offset

DFG difference frequency generation

e-e electron-electron

FWHM full width half maximum

IFR interface roughness

IV current-voltage

LLS lower laser state

LO longitudinal optical

MBE molecular beam epitaxy

mid-IR mid-infrared

MM metal-metal

MOCVD metal-organic chemical vapor deposition

NDC negative differential conductance

NDR negative differential resistance

NEGF non-equilibrium Green's function

QCL quantum cascade laser

SHG second harmonic generation

SI-SP semi-insulating surface plasmon

SL superlattice

SPPA single plasmon pole approximation

TDPT time-dependent perturbation theory

THz terahertz

ULS upper laser state

WS Wannier-Stark

References

[1] M. Weber, *Handbook of Laser Wavelengths*, Laser & Optical Science & Technology (CRC Press, 1998).

[2] G. Grynberg, A. Aspect, and C. Fabre, *Introduction to Quantum Optics* (Cambridge University Press, Cambridge, 2010).

[3] Pang Ka kit, (2014), (CC BY 2.5).

[4] J. Faist, *Quantum Cascade Lasers* (Oxford University Press, Oxford, 2013).

[5] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer, Berlin, 1999).

[6] H. Haug and S. Koch, *Quantum theory of the optical and electronic properties of semiconductors* (World Scientific, Singapore, 2004).

[7] J. H. Davies, *The Physics of Low-dimensional Semiconductors: An Introduction* (Cambridge university press, 1997).

[8] K. F. Renk, *Basics of Laser Physics* (Springer, Berlin Heidelberg, 2012).

[9] S. A. Ktitorov, G. S. Simin, and V. Y. Sindalovskii, "Bragg reflections and the high-frequency conductivity of an electronic solid-state plasma," Sov. Phys.–Sol. State **13**, 1872 (1972), [Fizika Tverdogo Tela **13**, 2230 (1971)].

[10] L. Esaki and R. Tsu, "Superlattice and negative differential conductivity in semiconductors," IBM J. Res. Dev. **14**, 61 (1970).

[11] A. Wacker, "Semiconductor superlattices: A model system for nonlinear transport," Phys. Rep. **357**, 1–111 (2002).

[12] A. A. Andronov, E. P. Dodin, D. I. Zinchenko, and Y. N. Nozdrin, "Towards Wannier-Stark THz superlattice laser," J. Phys.: Conf. Ser. **193**, 012079 (2009).

[13] A. Andronov, E. Dodin, D. Zinchenko, Y. Nozdrin, M. Ladugin, A. Marmalyuk, A. Padalitsa, V. Belyakov, I. Ladenkov, and A. Fefelov, "Stimulated emission at transitions between Wannier-Stark ladders in semiconductor superlattices," JETP Lett+ **102**, 207–211 (2015).

[14] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum Cascade Laser," Science **264**, 553–556 (1994).

[15] G. Scalari, M. I. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, "Broadband THz lasing from a photon-phonon quantum cascade structure," Opt. Express **18**, 8043 (2010).

[16] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, "Terahertz semiconductor-heterostructure laser," Nature **417**, 156 (2002).

[17] K. Unterrainer, R. Colombelli, C. Gmachl, F. Capasso, H. Y. Hwang, A. M. Sergent, D. L. Sivco, and A. Y. Cho, "Quantum cascade lasers with double metal-semiconductor waveguide resonators," Appl. Phys. Lett. **80**, 3060–3062 (2002).

[18] A. W. M. Lee, Q. Qin, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, "Real-time terahertz imaging over a standoff distance (> 25 meters)," Appl. Phys. Lett. **89**, 141125 (2006).

[19] P. H. Siegel *et al.*, "Terahertz technology," IEEE Trans. Microw. Theory **50**, 910–928 (2002).

[20] M. Tonouchi, "Cutting-edge terahertz technology," Nat. Photonics **1**, 97–105 (2007).

[21] M. A. Belkin, F. Capasso, F. Xie, A. Belyanin, M. Fischer, A. Wittmann, and J. Faist, "Room temperature terahertz quantum cascade laser source based on intracavity difference-frequency generation," Appl. Phys. Lett. **92**, 201101 (2008).

[22] M. A. Belkin and F. Capasso, "New frontiers in quantum cascade lasers: high performance room temperature terahertz sources," Phys. Scripta **90**, 118002 (2015).

[23] Q. Lu, D. Wu, S. Sengupta, S. Slivken, and M. Razeghi, "Room temperature continuous wave, monolithic tunable THz sources based on highly efficient mid-infrared quantum cascade lasers," Sci. Rep. **6**, 23595 (2016).

[24] C. Sirtori, P. Kruck, S. Barbieri, P. Collot, J. Nagle, M. Beck, J. Faist, and U. Oesterle, "GaAs/AlGaAs quantum cascade lasers," Appl. Phys. Lett. **73**, 3486 (1998).

[25] G. Strasser, S. Gianordoli, L. Hvozdara, W. Schrenk, K. Unterrainer, and E. Gornik, "GaAs/AlGaAs superlattice quantum cascade lasers at $\lambda \approx 13\,\mu$m," Appl. Phys. Lett. **75**, 1345 (1999).

[26] P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, J. B. Khurgin, Y. Dikmelik, X. Wang, J. Fan, and C. F. Gmachl, "Highly power-efficient quantum cascade lasers," Nat. Photonics **4**, 95–98 (2010).

[27] C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, "Recent progress in quantum cascade lasers and applications," Rep. Prog. Phys. **64**, 1533 (2001).

[28] R. F. Curl, F. Capasso, C. Gmachl, A. A. Kosterev, B. McManus, R. Lewicki, M. Pusharsky, G. Wysocki, and F. K. Tittel, "Quantum cascade lasers in chemical physics," Chem. Phys. Lett. **487**, 1–18 (2010).

[29] S. Fathololoumi, E. Dupont, C. Chan, Z. Wasilewski, S. Laframboise, D. Ban, A. Mátyás, C. Jirauschek, Q. Hu, and H. C. Liu, "Terahertz quantum cascade lasers operating up to $\sim$ 200 K with optimized oscillator strength and improved injection tunneling," Opt. Express **20**, 3866–3876 (2012).

[30] B. S. Williams, "Terahertz quantum-cascade lasers," Nat. Photonics **1**, 517 (2007).

[31] H. Richter, M. Wienold, L. Schrottke, K. Biermann, H. T. Grahn, and H. W. Hübers, "4.7-thz local oscillator for the great heterodyne spectrometer on sofia," IEEE Transactions on Terahertz Science and Technology **5**, 539–545 (2015).

[32] B. B. Hu and M. C. Nuss, "Imaging with terahertz waves," Opt. Lett. **20**, 1716–1718 (1995).

[33] D. M. Mittleman, R. H. Jacobsen, and M. C. Nuss, "T-ray imaging," IEEE J. Sel. Top. Quant. **2**, 679–692 (1996).

[34] S. Hunsche, M. Koch, I. Brener, and M. Nuss, "THz near-field imaging," Opt. Comm. **150**, 22–26 (1998).

[35] C. L. Koch-Dandolo, T. Filtenborg, K. Fukunaga, J. Skou-Hansen, and P. U. Jepsen, "Reflection terahertz time-domain imaging for analysis of an 18th century neoclassical easel painting," Appl. Opt. **54**, 5123–5129 (2015).

[36] C. Jirauschek and T. Kubis, "Modeling techniques for quantum cascade lasers," Appl. Phys. Rev. **1**, 011307 (2014).

[37] S. Khanal, L. Zhao, J. L. Reno, and S. Kumar, "Temperature performance of terahertz quantum-cascade lasers with resonant-phonon active-regions," J. Opt. **16**, 094001 (2014).

[38] D. Burghoff, T.-Y. Kao, N. Han, C. W. I. Chan, X. Cai, Y. Yang, D. J. Hayton, J.-R. Gao, J. L. Reno, and Q. Hu, "Terahertz laser frequency combs," Nat. Photonics **8**, 462–467 (2014).

[39] M. Rösch, G. Scalari, M. Beck, and J. Faist, "Octave-spanning semiconductor laser," Nat. Photonics **9**, 42–47 (2015).

[40] J. B. Khurgin, Y. Dikmelik, A. Hugi, and J. Faist, "Coherent frequency combs produced by self frequency modulation in quantum cascade lasers," Appl. Phys. Lett. **104**, 081118 (2014).

[41] S.-C. Lee, F. Banit, M. Woerner, and A. Wacker, "Quantum mechanical wavepacket transport in quantum cascade laser structures," Phys. Rev. B **73**, 245320 (2006).

[42] F. Capasso, C. Sirtori, and A. Y. Cho, "Coupled quantum well semiconductors with giant electric field tunable nonlinear optical properties in the infrared," IEEE J. Quantum Elect. **30**, 1313–1326 (1994).

[43] W. W. Chow and S. W. Koch, *Semiconductor - Laser Fundamentals* (Springer, Berlin, 1999).

[44] J. J. Sakurai, *Modern Quantum Mechanics*, 1st ed. (Addison Wesley, 1993).

[45] A. Wacker, "Quantum cascade laser: An emerging technology," in *Nonlinear Laser Dynamics*, edited by K. Lüdge (Wiley-VCH, Berlin, 2012).

[46] A. Hugi, G. Villares, S. Blaser, H. C. Liu, and J. Faist, "Mid-infrared frequency comb based on a quantum cascade laser," Nature **492**, 229 (2012).

[47] Y. Prior, "A complete expression for the third-order susceptibility $\chi^{(3)}$-perturbative and diagrammatic approaches," IEEE J. Quantum Elect. **20**, 37–42 (1984).

[48] T. Brandes, "Truncation method for green's functions in time-dependent fields," Phys. Rev. B **56**, 1213 (1997).

[49] B. J. Keay, S. Zeuner, S. J. Allen, K. D. Maranowski, A. C. Gossard, U. Bhattacharya, and M. J. W. Rodwell, "Dynamic localization, absolute negative conductance, and stimulated, multiphoton emission in sequential resonant tunneling semiconductor superlattices," Phys. Rev. Lett. **75**, 4102–4105 (1995).

[50] A. Wacker, A.-P. Jauho, S. Zeuner, and S. J. Allen, "Sequential tunneling in doped superlattices: Fingerprints of impurity bands and photon-assisted tunneling," Phys. Rev. B **56**, 13268 (1997).

[51] G. Platero and R. Aguado, "Sequential tunneling current through semiconductor superlattices under intense THz radiation," Appl. Phys. Lett. **70**, 3546–3548 (1997).

[52] F. Capasso, J. Faist, and C. Sirtori, "Mesoscopic phenomena in semiconductor nanostructures by quantum design," J. Math. Phys. **37**, 4775 (1996).

[53] F. Compagnone, A. Di Carlo, and P. Lugli, "Monte Carlo simulation of electron dynamics in superlattice quantum cascade lasers," Appl. Phys. Lett. **80**, 920 (2002).

[54] D. Indjin, P. Harrison, R. W. Kelsall, and Z. Ikonic, "Self-consistent scattering theory of transport and output characteristics of quantum cascade lasers," J. Appl. Phys. **91**, 9019 (2002).

[55] D. Paulavičius, V. Mitin, and M. A. Stroscio, "Hot-optical-phonon effects on electron relaxation in an AlGaAs/GaAs quantum cascade laser structure," J. Appl. Phys. **84**, 3459 (1998).

[56] S. Slivken, V. I. Litvinov, M. Razeghi, and J. R. Meyer, "Relaxation kinetics in quantum cascade lasers," J. Appl. Phys. **85**, 665 (1999).

[57] C. Becker, C. Sirtori, H. Page, A. Robertson, V. Ortiz, and X. Marcadet, "Influence of confined phonon modes on the thermal behavior of AlAs/GaAs quantum cascade structures," Phys. Rev. B **65**, 085305 (2002).

[58] P. Hyldgaard and J. W. Wilkins, "Electron-electron scattering in far-infrared quantum cascade lasers," Phys. Rev. B **53**, 6889 (1996).

[59] P. Harrison, "The nature of the electron distribution functions in quantum cascade lasers," Appl. Phys. Lett. **75**, 2800 (1999).

[60] S. Tortora, F. Compagnone, A. Di Carlo, P. Lugli, M. T. Pellegrini, M. Troccoli, and G. Scamarcio, "Theoretical study and simulation of electron dynamics in quantum cascade lasers," Physica B **272**, 219 (1999).

[61] R. C. Iotti and F. Rossi, "Carrier thermalization versus phonon-assisted relaxation in quantum cascade lasers: A monte carlo approach," Appl. Phys. Lett. **78**, 2902–2904 (2001).

[62] R. C. Iotti and F. Rossi, "Nature of charge transport in quantum-cascade lasers," Phys. Rev. Lett. **87**, 146603 (2001).

[63] H. Callebaut, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, "Importance of electron-impurity scattering for electron transport in terahertz quantum-cascade lasers," Appl. Phys. Lett. **84**, 645 (2004).

[64] O. Bonno, J. Thobel, and F. Dessenne, "Modeling of electron-electron scattering in monte carlo simulation of quantum cascade lasers," J. Appl. Phys. **97**, 043702 (2005).

[65] C. Jirauschek and P. Lugli, "Monte-carlo-based spectral gain analysis for terahertz quantum cascade lasers," J. Appl. Phys. **105**, 123102 (2009).

[66] H. Callebaut and Q. Hu, "Importance of coherence for electron transport in terahertz quantum cascade lasers," J. Appl. Phys. **98**, 104505 (2005).

[67] S. Kumar and Q. Hu, "Coherence of resonant-tunneling transport in terahertz quantum-cascade lasers," Phys. Rev. B **80**, 245316 (2009).

[68] A. Gordon and D. Majer, "Coherent transport in semiconductor heterostructures: A phenomenological approach," Phys. Rev. B **80**, 195317 (2009).

[69] E. Dupont, S. Fathololoumi, and H. C. Liu, "Simplified density-matrix model applied to three-well terahertz quantum cascade lasers," Phys. Rev. B **81**, 205311 (2010).

[70] R. Terazzi and J. Faist, "A density matrix model of transport and radiation in quantum cascade lasers," New J. Phys. **12**, 033045 (2010).

[71] I. Waldmueller, W. W. Chow, E. W. Young, and M. C. Wanke, "Nonequilibrium many-body theory of intersubband lasers," IEEE J. Quantum Elect. **42**, 292 (2006).

[72] C. Weber, A. Wacker, and A. Knorr, "Density-matrix theory of the optical dynamics and transport in quantum cascade structures: The role of coherence," Phys. Rev. B **79**, 165322 (2009).

[73] A. Wacker, "Gain in quantum cascade lasers and superlattices: A quantum transport theory," Phys. Rev. B **66**, 085326 (2002).

[74] T. Schmielau and M. Pereira, "Nonequilibrium many body theory for quantum transport in terahertz quantum cascade lasers," Appl. Phys. Lett. **95**, 231111 (2009).

[75] T. Kubis, C. Yeh, P. Vogl, A. Benz, G. Fasching, and C. Deutsch, "Theory of nonequilibrium quantum transport and energy dissipation in terahertz quantum cascade lasers," Phys. Rev. B **79**, 195323 (2009).

[76] G. Haldaś and, A. Kolek, and I. Tralle, "Modeling of mid-infrared quantum cascade laser by means of nonequilibrium green's functions," IEEE J. Quantum Elect. **47**, 878 (2011).

[77] J. B. Khurgin, "Inhomogeneous origin of the interface roughness broadening of intersubband transitions," Appl. Phys. Lett. **93**, 091104 (2008).

[78] J. B. Khurgin, Y. Dikmelik, P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, and C. F. Gmachl, "Role of interface roughness in the transport and lasing characteristics of quantum-cascade lasers," Appl. Phys. Lett. **94**, 091101 (2009).

[79] G. H. Wannier, "The structure of electronic excitation levels in insulating crystals," Phys. Rev. **52**, 191 (1937).

[80] A. Bruno-Alfonso and D. R. Nacbar, "Wannier functions of isolated bands in one-dimensional crystals," Phys. Rev. B **75**, 115428 (2007).

[81] M. Franckié, *Modeling Quantum Cascade Lasers*, Ph.D. thesis, Lund University, Department of Physics (2016).

[82] G. D. Mahan, *Many-Particle Physics* (Plenum, New York, 2000).

[83] D. C. Langreth, "Linear and nonlinear response theory with applications," in *Linear and Nonlinear Electron Transport in Solids*, edited by J. T. Devreese and V. E. van Doren (Plenum Press, New York, 1976).

[84] D. Binosi and L. Theußl, "Jaxodraw: A graphical user interface for drawing feynman diagrams," Comput. Phys. Commun. **161**, 76 – 86 (2004).

[85] H. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 1996).

[86] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-particle Systems* (Dover Publications, 2003).

[87] S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, "1.9 THz quantum-cascade lasers with one-well injector," Appl. Phys. Lett. **88**, 121123 (2006).

[88] M. S. Vitiello, G. Scamarcio, J. Faist, G. Scalari, C. Walther, H. E. Beere, and D. A. Ritchie, "Probing quantum efficiency by laser-induced hot-electron cooling," Appl. Phys. Lett. **94**, 021115 (2009).

[89] J. H. Shirley, "Solution of the Schrödinger equation with a Hamiltonian periodic in time," Phys. Rev. **138**, B979–B987 (1965).

[90] L. Hedin, "New method for calculating the one-particle Green's function with application to the electron-gas problem," Phys. Rev. **139**, A796 (1965).

[91] R. Nelander, *Lineshape in Quantum Cascade Lasers - Temperature, Screening and Broadening*, Ph.D. thesis, Lund University (2009).

[92] M. I. Amanti, G. Scalari, R. Terazzi, M. Fischer, M. Beck, J. Faist, A. Rudra, P. Gallo, and E. Kapon, "Bound-to-continuum terahertz quantum cascade laser with a single-quantum-well phonon extraction/injection stage," New J. Phys. **11**, 125022 (2009).

[93] D. O. Winge, M. Lindskog, and A. Wacker, "Nonlinear response of quantum cascade structures," Appl. Phys. Lett. **101**, 211113 (2012).

[94] L. Li, L. Chen, J. Zhu, J. Freeman, P. Dean, A. Valavanis, A. G. Davies, and E. H. Linfield, "Terahertz quantum cascade lasers with >1 W output powers," Electron. Lett. **50**, 309–311 (2014).

[95] R. Nelander and A. Wacker, "Temperature dependence of the gain profile for terahertz quantum cascade lasers," Appl. Phys. Lett. **92**, 081102 (2008).

[96] M. S. Vitiello, R. C. Iotti, F. Rossi, L. Mahler, A. Tredicucci, H. E. Beere, D. A. Ritchie, Q. Hu, and G. Scamarcio, "Non-equilibrium longitudinal and transverse optical phonons in terahertz quantum cascade lasers," Appl. Phys. Lett. **100**, 091101 (2012).

[97] Y. B. Shi and I. Knezevic, "Nonequilibrium phonon effects in midinfrared quantum cascade lasers," J. Appl. Phys. **116**, 123105 (2014).

[98] D. Burghoff, T.-Y. Kao, D. Ban, A. W. M. Lee, Q. Hu, and J. Reno, "A terahertz pulse emitter monolithically integrated with a quantum cascade laser," Appl. Phys. Lett. **98**, 061112 (2011).

[99] B. Williams, S. Kumar, Q. Hu, and J. Reno, "Resonant-phonon terahertz quantum-cascade laser operating at 2.1 THz ($\lambda \sim 141$ μm)," Electron. Lett. **40**, 431–433 (2004).

[100] A. Benz, G. Fasching, A. M. Andrews, M. Martl, K. Unterrainer, T. Roch, W. Schrenk, S. Golka, and G. Strasser, "Influence of doping on the performance of terahertz quantum-cascade lasers," Appl. Phys. Lett. **90**, 101107 (2007).

[101] F. Banit, S.-C. Lee, A. Knorr, and A. Wacker, "Self-consistent theory of the gain linewidth for quantum cascade lasers," Appl. Phys. Lett. **86**, 041108 (2005).

[102] C. A. Balanis, *Antenna Theory: Analysis and Design*, 2nd ed. (Wiley, New York, 1997).

[103] S. Kohen, B. Williams, and Q. Hu, "Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators," J. Appl. Phys. **97**, 053106 (2005).

[104] M. Franckié, D. O. Winge, J. Wolf, V. Liverini, E. Dupont, V. Trinité, J. Faist, and A. Wacker, "Impact of interface roughness distributions on the operation of quantum cascade lasers," Opt. Express **23**, 5201–5212 (2015).

[105] K. A. Krivas, D. O. Winge, M. Franckié, and A. Wacker, "Influence of interface roughness in quantum cascade lasers," J. Appl. Phys. **118**, 114501 (2015).

[106] L. Ajili, G. Scalari, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies, "High power quantum cascade lasers operating at $\lambda \simeq 87$ and 130 μm," Appl. Phys. Lett. **85**, 3986–3988 (2004).

[107] M. S. Vitiello, G. Scamarcio, V. Spagnolo, S. S. Dhillon, and C. Sirtori, "Terahertz quantum cascade lasers with large wall-plug efficiency," Appl. Phys. Lett. **90**, 191115 (2007).

[108] M. A. Belkin, J. A. Fan, S. Hormoz, F. Capasso, S. P. Khanna, M. Lachab, A. G. Davies, and E. H. Linfield, "Terahertz quantum cascade lasers with copper metal-metal waveguides operating up to 178 k," Opt. Express **16**, 3242–3248 (2008).

[109] J. A. Fan, M. A. Belkin, F. Capasso, S. P. Khanna, M. Lachab, A. G. Davies, and E. H. Linfield, "Wide-ridge metal-metal terahertz quantum cascade lasers with high-order lateral mode suppression," Appl. Phys. Lett. **92**, 031106 (2008).

[110] C. Gmachl, A. Belyanin, D. L. Sivco, M. L. Peabody, N. Owschimikow, A. M. Sergent, F. Capasso, and A. Y. Cho, "Optimized second-harmonic generation in quantum cascade lasers," IEEE J. Quantum Elect. **39**, 1345–1355 (2003).

[111] C. Sirtori, F. Capasso, D. L. Sivco, S. Chu, and A. Y. Cho, "Observation of large second order susceptibility via intersubband transitions at $\lambda \sim 10$ μm

in asymmetric coupled AlInAs/GaInAs quantum wells," Appl. Phys. Lett. **59**, 2302–2304 (1991).

[112] N. Owschimikow, C. Gmachl, A. Belyanin, V. Kocharovsky, D. L. Sivco, R. Colombelli, F. Capasso, and A. Y. Cho, "Resonant second-order nonlinear optical processes in quantum cascade lasers," Phys. Rev. Lett. **90**, 043902 (2003).

[113] A. Wacker, "Coherence and spatial resolution of transport in quantum cascade lasers," phys. stat. sol. (c) **5**, 215 (2008).

[114] M. Lindskog, J. M. Wolf, V. Trinite, V. Liverini, J. Faist, G. Maisons, M. Carras, R. Aidam, R. Ostendorf, and A. Wacker, "Comparative analysis of quantum cascade laser modeling based on density matrices and nonequilibrium green's functions," Appl. Phys. Lett. **105**, 103106 (2014).

[115] S. Kumar, Q. Hu, and J. L. Reno, "186 K operation of terahertz quantum-cascade lasers based on a diagonal design," Appl. Phys. Lett. **94**, 131105 (2009).

[116] X. Gao, D. Botez, and I. Knezevic, "X-valley leakage in GaAs/AlGaAs quantum cascade lasers," Appl. Phys. Lett. **89**, 191119 (2006).

[117] B. A. Burnett and B. S. Williams, "Density matrix model for polarons in a terahertz quantum dot cascade laser," Phys. Rev. B **90**, 155309 (2014).

[118] W. Freeman, "Self-consistent calculation of dephasing in quantum cascade structures within a density matrix method," Phys. Rev. B **93**, 205301 (2016).

Part II

The papers

Paper I

Paper II

Paper III

Paper IV

Paper V

Paper VI