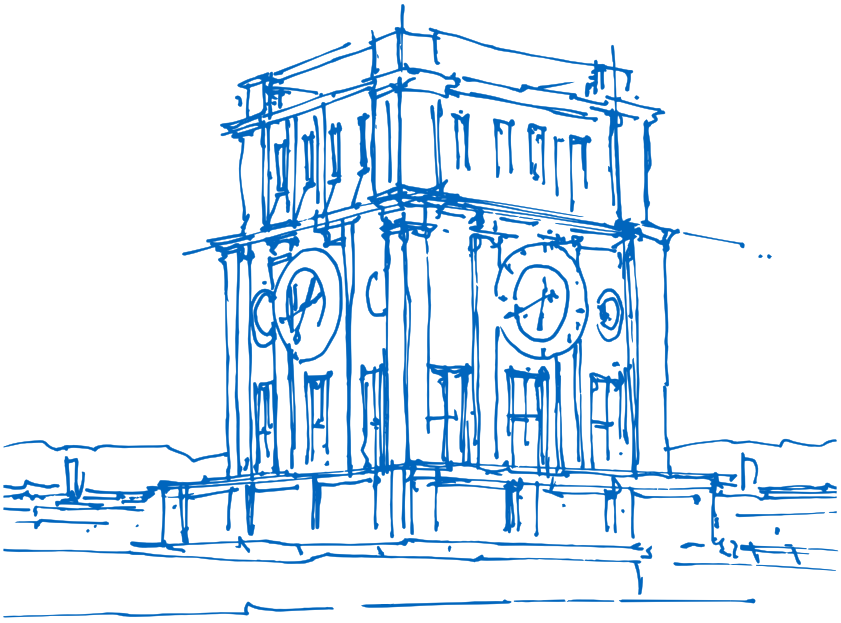
On the Modeling and Simulation of the Dynamical Behavior of Quantum Cascade Lasers

Michael Riesch



TUM School of Computation, Information and Technology
Technische Universität München

On the Modeling and Simulation of the Dynamical Behavior of Quantum Cascade Lasers

Michael Riesch

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:
Prof. Dr.-Ing. Norbert Hanik

Prüfer*innen der Dissertation:
1. Prof. Dr.-Ing. Christian Jirauschek
2. Prof. Dr. Kai Müller

Die Dissertation wurde am 17.06.2022 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 21.04.2023 angenommen.

To my parents Doris and Herbert,
and to Andrea.

Abstract

I think this is a song of hope!
— Robert Plant, The Song Remains the Same

Quantum cascade lasers (QCLs) are relatively novel sources of mid-infrared and terahertz radiation and thus promising candidates to close the so-called terahertz gap. In their envisaged applications advanced operating modes, such as frequency comb operation or ultra-short pulse generation, are desired. Those advanced operating modes are significantly more complex than plain continuous wave operation and require a clear understanding of the processes in QCLs. This understanding can be gained by modeling and simulation of their dynamical behaviour. The thesis at hand provides a solid basis for the upcoming endeavors in the field by reviewing the theoretical modeling of the QCL dynamics and presenting a flexible software framework for the computer-aided simulation of those models.

The thesis begins with a compact description of the required theoretical basics to make the thesis self-contained. Then, the model of the QCL dynamics is developed step-by-step from the very foundations of the underlying theory. The flexible open-source software framework is presented as a next step. It implements several best practices to make it a sustainable scientific software project, which are discussed in detail. A detailed description of the numerical methods and the parallelization techniques used in the software framework follows. Finally, the software framework is applied to actual problems in the field. Thereby, it is verified that the framework is already able to answer questions raised by experimental findings.

Kurzfassung (German)

Quantenkaskadenlaser sind relativ neuartige Quellen von mittlerer Infrarot- und Terahertz-Strahlung und sind daher vielversprechende Kandidaten, um die sogenannte Terahertz-Lücke zu schließen. In den ihnen angedachten Anwendungen sind fortgeschrittene Betriebsmodi (wie etwa Frequenzkammbetrieb oder die Generation von ultra-kurzen Pulsen) erforderlich. Diese fortgeschrittenen Betriebsmodi sind wesentlich komplexer als der simple Dauerstrichbetrieb und erfordern ein klares Verständnis der Prozesse in den Quantenkaskadenlasern. Dieses Verständnis kann durch Modellierung und Simulation ihres dynamischen Verhaltens erworben werden. Die vorliegende Arbeit stellt eine solide Basis für die bevorstehenden Unterfangen in diesem Gebiet zur Verfügung, indem sie eine Übersicht über die theoretische Modellierung der Dynamik der Quantenkaskadenlaser bietet und ein flexibles Software-Framework für die computerunterstützte Simulation der Modelle präsentiert.

Die Arbeit beginnt mit einer kompakten Beschreibung der benötigten theoretischen Grundlagen, um die Arbeit in sich abgeschlossen zu gestalten. Danach wird das dynamische Modell von Quantenkaskadenlasern Schritt für Schritt von den theoretischen Grundlagen hergleitet. Das flexible open-source Software-Framework wird als nächster Schritt präsentiert. Es implementiert viele Maßnahmen aus der Softwareentwicklung, die es zu einem nachhaltigen Softwareprojekt machen und die im Detail besprochen werden. Eine detaillierte Beschreibung der verwendeten numerischen Methoden und der Parallelisierung von Berechnungen folgt anschließend. Schließlich wird das Framework auf tatsächliche Probleme aus dem Fachgebiet angewendet. Damit wird demonstriert, dass das Framework bereits jetzt in der Lage ist, Fragen zu beantworten, die durch experimentelle Erkenntnisse aufgeworfen wurden.

Contents

Abstract vii

Kurzfassung (German) ix

Nomenclature xiii

1 Introduction 1
- 1.1 Lasers . . . 2
- 1.2 Frequency combs and ultra-short pulse generation . . . 3
- 1.3 Quantum cascade lasers . . . 5
- 1.4 Overview of this work . . . 7

2 Theoretical background 9
- 2.1 Classical electrodynamics . . . 9
- 2.2 Fundamentals of quantum mechanics . . . 13
- 2.3 Basic theory of semiconductors and semiconductor heterostructures . . . 21
- 2.4 Summary . . . 26

3 Modeling the dynamics of quantum cascade lasers 27
- 3.1 Electron dynamics in the gain medium . . . 27
- 3.2 Electromagnetic field in the resonator cavity . . . 35
- 3.3 Overview of the coupled model and simulation software . . . 39
- 3.4 Summary . . . 41

4 An open-source solver for the Maxwell-Bloch equations 43
- 4.1 The Maxwell-Bloch equations and existing solver tools . . . 44
- 4.2 Implementation of mbsolve . . . 46
- 4.3 Application and simulation examples . . . 51
- 4.4 Summary . . . 57

5 Ensuring quality in scientific software engineering 59
- 5.1 Best practices in scientific software engineering . . . 60
- 5.2 Implementation of the project skeleton . . . 63
- 5.3 Creating a skeleton instance . . . 67
- 5.4 Summary . . . 68

6 Numerical treatment of the Maxwell-Bloch equations 69
- 6.1 Numerical methods for Maxwell's equations . . . 69
- 6.2 Completely positive trace-preserving methods for the master equation . . . 72
- 6.3 Implementation and single-thread performance comparison . . . 76
- 6.4 Summary . . . 83

7 Solving the Maxwell-Bloch equations on parallel architectures 85
- 7.1 Parallel implementation of the FDTD method . . . 86
- 7.2 Advanced FDTD implementation . . . 87
- 7.3 Performance measurements . . . 89

7.4 Summary 91

8 Application of the simulation framework to different problems 93
8.1 Colliding pulse mode-locking in quantum cascade lasers 93
8.2 Modeling harmonic mode-locking in quantum cascade lasers 95
8.3 Dynamics of a quantum cascade laser with a graphene reflector 98
8.4 Summary 99

9 Conclusion and outlook 101

Bibliography 103

Publications 115

Acknowledgments 119

Curriculum vitae 121

Nomenclature

Habib: Give a dog a bad name!
Gladstone: We gave our dog a bad name. Colin. Terrible name! You have no idea how many people are called Colin.
— Maggie Habib and Frank Gladstone, The Thin Blue Line

$*$	Convolution operator
$[\cdot,\cdot]$	Commutator, $[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}$
$\langle\cdot\vert\cdot\rangle$	Inner product
$\langle\cdot\rangle$	Expectation value
$\langle\cdot\rangle_{3D}$	Macroscopic expectation value
$\mathcal{F}\{\cdot\}$	Fourier transform
$\mathcal{F}^{-1}\{\cdot\}$	Inverse Fourier transform
$\mathcal{O}(\cdot)$	Landau notation for the asymptotic upper bound
∇	Differential operator nabla, $\nabla = [\partial_x, \partial_y, \partial_z]$
$\otimes$	Tensor product
$\mathrm{Tr}\{\cdot\}$	Trace operation
$\mathrm{Tr}_A\{\cdot\}$	Partial trace operation over A
$\underline{A}$	Complex number
$\underline{A}^*$	Complex conjugate of A
$\hat{A}$	Quantum mechanical operator
$\hat{A}^\dagger$	Hermitian conjugate of $\hat{A}$
$\boldsymbol{A}$	Electromagnetic vector potential
A_{qm}	Cross section area of the active region
$\boldsymbol{B}, B$	Magnetic field
C	Courant number
c	Velocity of light in vacuum
$\mathcal{D}$	Dissipation superoperator
$\boldsymbol{D}$	Displacement field
$\boldsymbol{E}, E$	Electric field
e	Elementary charge
$\boldsymbol{F}$	Force field
f_{rt}	Round trip frequency
$\hat{H}$	Hamiltonian operator
$\hbar$	Reduced Planck constant
$\mathcal{H}$	Hamiltonian function
$\boldsymbol{H}, H$	Magnetic H-field
i	Imaginary unit
$\boldsymbol{J}, J$	Current density
$\boldsymbol{k}$	Wave vector
$\mathcal{L}$	Liouvillian superoperator
L_{cav}	Cavity length
$\boldsymbol{M}, M$	Magnetization field
m	Mass
m^*	Effective mass in growth direction
$m^{\vert\vert}$	Effective in-plane mass
m_0	Electron mass
N	Number of relevant energy levels
n	Refractive index
n_{3D}	Density of particles in 3D
$\boldsymbol{p}$	Momentum
$\boldsymbol{P}, P$	Polarization field
q	Electric charge
$\boldsymbol{r}$	Position vector $\boldsymbol{r} = [x, y, z]^T$
S	In-plane cross section area

Δt Temporal discretization size

t Time

T_{rt} Round trip time

$\mathcal{U}$ Unitary evolution superoperator

ΔV Test volume

$\mathcal{V}$ Evolution superoperator of the reduced system

$\boldsymbol{v}$ Velocity

V Energy potential

V_{c} Conduction band profile

Δx Spatial discretization size

x Coordinate in light propagation direction

y Transversal coordinate, perpendicular to light propagation and growth direction

z Coordinate in growth direction of heterostructure

α_0 Linear loss term

$\underline{\beta}$ Complex propagation constant

Γ Overlap factor

γ Scattering or dephasing rate

$\delta(x)$ Dirac distribution

δ_{ij} Kronecker delta

ϵ Energy

ϵ_i Eigenenergy of the state $|i\rangle$

ε_0 Permittivity of vacuum

ε_{r} Relative permittivity

λ Wavelength

μ_0 Permeability of vacuum

μ_{r} Relative permeability

$\hat{\boldsymbol{\mu}}, \hat{\mu}$ Dipole moment operator

$\hat{\rho}$ Density operator

ρ Charge density

σ Electric conductivity

τ^{-1} Inverse population lifetime

ϕ Basis function

φ Electromagnetic scalar potential

χ_{e} Electric susceptibility

χ_{m} Magnetic susceptibility

$|\Psi\rangle$ Quantum mechanical state in Dirac notation

Ψ Wave function

ψ Wave function envelope

ω Angular frequency

1 Introduction

> Way out west there was this fellow... fellow I wanna tell you about. Fellow by the name of Jeff Lebowski. At least that was the handle his loving parents gave him, but he never had much use for it himself.
>
> — *The Stranger, The Big Lebowski*

> Space. It seems to go on and on forever. But then you get to the end and the gorilla starts throwing barrels at you.
>
> — *Philip J. Fry, Futurama*

Since their theoretical description by Kazarinov and Suris [1] in 1971 and their first experimental realization by Faist et al. [2] in 1994, quantum cascade lasers (QCLs) have attracted considerable attention. In contrast to conventional laser diodes – where the radiation frequency is largely determined by the semiconductor band gap of the active region – the active region of the QCL can be engineered to emit radiation in the mid-infrared and terahertz (far-infrared) regime [3]. This portion of the electromagnetic frequency spectrum lies between the realms of optics and microwaves, and features many interesting applications in spectroscopy and imaging since it contains the fundamental vibrational bands of a large number of molecular compounds, vapours and gases [4], [5]. Due to the lack of powerful and compact sources, however, the huge potential of this frequency regime could not be exploited for a long time. For example, the lack of practical technologies for the terahertz regime (often referred to as terahertz gap) has impeded the development of novel measurement techniques to be applied in astrophysical observation, environmental monitoring, the detection of explosives and illegal drugs, and terahertz spectroscopy of DNA [5], [6]. Devices such as terahertz QCLs, which were realized for the first time by Köhler et al. [7] in 2002, aim to close this gap. This fact makes this relatively novel laser type attractive for a variety of applications in several fields, such as physics, chemistry, biology, medicine, astronomy and public safety [8].

Over the years, the performance of QCLs with respect to operating temperature, output power, and the available frequency range has improved considerably [3], [5]. The development and optimization of the QCL designs has gone hand in hand with detailed modeling and computer-aided simulations [9], which underlines the importance of computational science as third pillar of science next to theory and experiment. However, the application range of QCLs does not depend on the performance alone, but also on the availability of different operating modes. For example, for applications in spectroscopy the QCLs have to be driven in frequency-comb operation or are supposed to emit ultra-short pulses [8]. Those envisaged operating modes are significantly more complex than plain continuous wave operation and require a clear understanding of the processes in QCLs that define their dynamical behavior. Similar to the design optimization, this understanding can only be gained by combining experimental and theoretical work, where the latter is divided into the modeling of the underlying physical processes and accurate computer simulations. The thesis at hand aims to provide a solid basis for the upcoming endeavors in this field by reviewing the theoretical modeling of the QCL dynamics and presenting a flexible software framework for the computer-aided simulation of those models.

In this chapter, we start with a brief introduction of the basic operating principle of lasers in Section 1.1 and their advanced operating modes, such as frequency comb operation and ultra-short pulse generation, in Section 1.2. Then, Section 1.3 discusses the basic operating principle of quantum cascade lasers and their different subtypes. After those introductory sections, which should make the reader acquainted with the devices to be modeled and simulated as well as with the motivation to do so, this chapter is concluded by an overview of the thesis at hand.

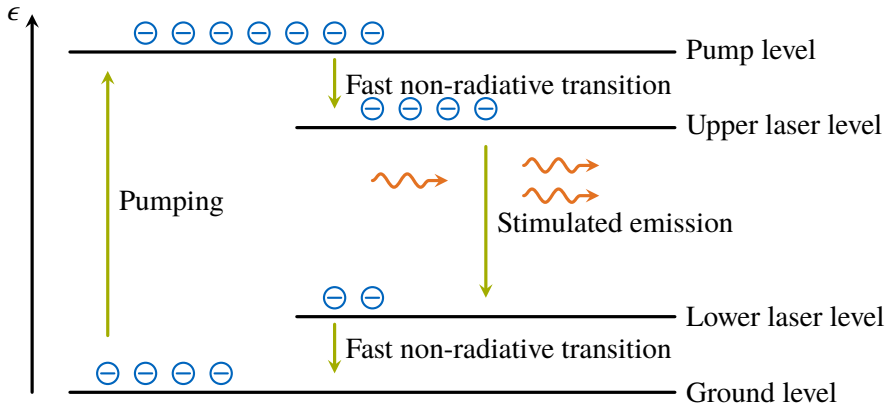


Figure 1.1 Example of a four-level laser pumping scheme. An external source of energy pumps electrons from the ground level to the pumping level. Fast non-radiative transitions ensure that there are many electrons in the upper laser level and only few electrons in the lower laser level (population inversion). Between the upper and lower laser level, on the other hand, the non-radiative transitions must be significantly slower. Then, stimulated emission dominates the transitions from the upper to the lower laser level and exact copies of the incoming photons are emitted.

1.1 Lasers

Lasers come in different forms and are used in a multitude of applications, making them one of the most influential inventions of the twentieth century. Their basic operating principle is embedded in their name since "laser" stands for "light amplification by stimulated emission of radiation". In this section, we describe this operating principle briefly. For a more thorough coverage of the topic the reader is referred to the comprehensive text book by Siegman [10], which is used as basis for the following introduction.

Stimulated emission is an effect that appears during the interaction of electrons and photons. To explain it qualitatively, let us assume that electrons can occupy different discrete energy levels (states) in a system. The lowest energy level is usually referred to as the ground state. If additional energy is supplied to the system, the electrons can be brought into an exited state. An electron in an excited state will eventually return to a lower energy level (if available), causing the difference in energy $\Delta\epsilon$ to be emitted. Such emission processes are grouped into radiative (emission of radiation) and non-radiative processes depending on the form in which the energy is emitted, and into spontaneous and stimulated emission. Spontaneous emission of radiation occurs without any external trigger and is a random process. As a consequence, the properties of the emitted radiation, such as frequency and phase, are purely random. Stimulated emission of radiation, on the other hand, is triggered by an incoming photon (the smallest quantum of radiation) and generates a second photon during the transition of the electron from the upper energy level to the lower energy level. Here, the second photon is an exact copy of the incoming photon (with respect to frequency, phase, polarization and direction). As a consequence, the incoming radiation is amplified and the resulting radiation is highly coherent, where coherence denotes the constant phase relationship between all the photons. Figure 1.1 depicts this effect schematically.

The medium in which the light amplification process takes place is usually referred to as the gain medium. In real-life lasers, a multitude of materials and material systems in gaseous (e.g., helium-neon laser, carbon dioxide laser), liquid (e.g., dye lasers), and solid (e.g., titanium-sapphire laser, semiconductor lasers) form is used as gain medium. All of them have in common that the electrons in the medium can only assume discrete energy values. For example, in a noble gas laser the energy values are defined by the atomic orbitals of the gas atoms. In semiconductor lasers, on the other hand, usually the band gap leads to a discretization of the allowed energy values. At this point, it should be noted that stimulated emission of radiation only occurs if the energy difference $\Delta\epsilon$ between the upper and lower energy level matches the photon energy ϵ given by Planck's relation

$$\epsilon = \hbar\omega, \tag{1.1}$$

where $\hbar$ is the reduced Planck's constant and ω is the angular frequency of the photon. This fact establishes a relation between the gain spectrum (the spectrum of radiation that a certain laser amplifies) and the material properties of its gain medium. In order to amplify light of a certain frequency, a gain medium with suitable properties (e.g., the atomic orbitals of a noble gas, or the band gap of a semiconductor) has to be selected.

Furthermore, we note that the gain medium is only functional if there are many electrons in the upper energy level (upper laser level, ULL) and only few electrons in the lower energy level (lower laser level, LLL). This state is called population inversion and is achieved with the help of a pumping mechanism. There are different types of pumping mechanisms that can be grouped by the energy form in use (e.g., optical pumping, electrical pumping, ...). Although two levels are sufficient for stimulated emission, at least one additional energy level is required to achieve population inversion. However, pumping schemes with more energy levels are commonly used in real-life lasers. Here, the pumping mechanism and non-radiative processes cause electron transitions into the ULL and out of the LLL, while the transition between the ULL and the LLL is of course a dominantly radiative one. An exemplary four-level pumping scheme is depicted in Fig. 1.1.

Amplifiers with a gain medium alone are widely used in fiber telecommunications, where the incoming signal has to be amplified to bridge long distances. Should the laser act as oscillator (source), however, a feedback mechanism is required. This mechanism is usually provided by a suitable resonator cavity, which may come in different forms. For example, in a ring cavity the radiation is guided in a closed loop, so that the amplified radiation is fed back to the input. At some point, a portion of the radiation is outcoupled. This portion is the actual product of the laser oscillator. Another significant type of resonator cavity is the Fabry-Pérot cavity, in which the radiation is reflected between two mirrors and amplified by the gain medium between the mirrors. In this case, outcoupling can be achieved by making one of the mirrors semi-transparent. It should be noted that a small amount of radiation is required in order to start the laser up. While this can be achieved by using a different laser (laser seeding), often it is sufficient to exploit spontaneous emission, which is beneficial for lasing operation initially (after the laser has started up, it impedes the coherence of the radiation).

In summary, every laser needs a gain medium, in which the electrons can only assume discrete energy values and a pumping mechanism that creates population inversion (a state in which more electrons are in excited states rather than in the ground state). Laser sources additionally requires a resonator cavity that provides a feedback mechanism. Let us keep these basic principles in mind while we turn our attention towards the different operating modes of a laser.

1.2 Frequency combs and ultra-short pulse generation

The operating modes of a laser are distinguished by its main product, namely the emitted electromagnetic radiation. If the power of the radiation is (quasi-)constant over time, the laser is said to be in (quasi-)continuous wave (cw) operation. On the other hand, if the power varies significantly so that the radiation practically vanishes for a certain period of time, the laser is operated in pulsed mode. Both operating modes have their applications and different measures may be required to operate a laser in a certain mode.

In order to describe the operating modes, the time profile of the emitted radiation as well as its spectrum in frequency domain have to be evaluated. In the easiest case, the spectrum consists of a single sharp peak at a certain frequency. For the sake of simplicity we neglect the line shape of the peak for now and assume that it is infinitesimally narrow. Then, we recall that the gain spectrum $G(\omega)$ of a laser depends heavily on frequency and deduce that if a certain frequency is emitted then the gain for this frequency is sufficiently large to overcome all losses in the laser (usually called the lasing threshold). Also, the frequency must belong to the set of allowed frequencies of the resonator cavity. Those frequencies are commonly referred to as modes, and if the laser only produces a certain frequency it is said to be in single-mode operation. This operating mode might be considered the most trivial one from a theoretical point of view, but in practice it may be difficult to drive a laser into single-mode operation. In fact, the gain spectrum often allows the amplification of multiple modes, which have to be suppressed if single-mode operation is desired.

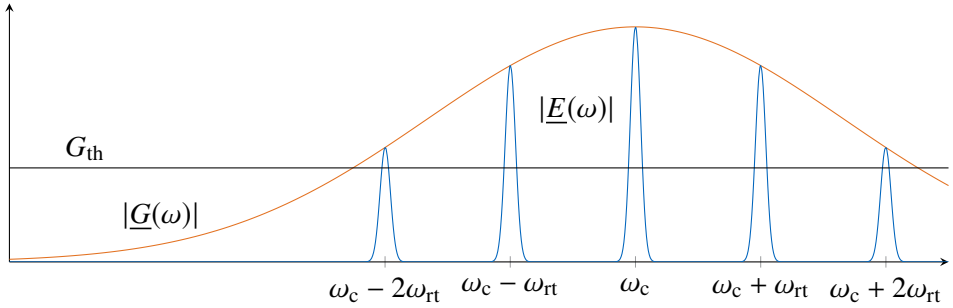


Figure 1.2 Exemplary spectrum of the radiation emitted by a Fabry-Pérot laser. The different modes are centered around the carrier frequency ω_{c} and are separated by the round trip frequency $\omega_{\mathrm{rt}} = 2\pi f_{\mathrm{rt}}$. The spectrum only contains the modes that experience a sufficiently large gain, i.e., the gain spectrum $|\underline{G}(\omega)|$ at the corresponding frequency is larger than the threshold value G_{th}.

Many applications, however, require multi-mode operation. In this case, the spectrum of the emitted radiation resembles a comb and coined the name frequency comb [11]. As an example, Fig. 1.2 depicts qualitatively the spectrum of the radiation emitted by a Fabry-Pérot laser. In a Fabry-Pérot cavity the electromagnetic field can only exist in the form of standing waves due to the imposed boundary conditions. The problem can be directly related to the standard physics exercise of a string with two fixed ends [12] and yields

$$\omega_k = 2\pi k \frac{c}{2L_{\mathrm{cav}} n} = 2\pi k \frac{1}{T_{\mathrm{rt}}} = 2\pi k f_{\mathrm{rt}} \tag{1.2}$$

for the allowed frequencies ω_k, where c is the velocity of light in vacuum, n is the refractive index of the medium in the cavity, and L_{cav} is the length of the cavity. The round trip frequency f_{rt} is commonly introduced for brevity. It constitutes the inverse of the round trip time $T_{\mathrm{rt}} = 2L_{\mathrm{cav}} n/c$, which is the time that the radiation requires for one round trip through the resonator cavity. Therefore, the spectrum in Fig. 1.2 consists of multiple modes that are separated by the round trip frequency [10]. Since the round trip frequency is typically in the gigahertz regime but the frequencies ω_k are in the optical domain, frequency combs provide a link between optics and microwaves. In essence, the link is established with the help of a nonlinear mixing process, which causes two adjacent modes to produce a beating signal similar to two tuning forks with slightly different tuning. This approach is heavily used in high-precision optical metrology, e.g., to read out the information in the optical signal using state-of-the-art electronics [11], [13].

Frequently, frequency combs are used to generate ultra-short pulses, i.e., pulses that cannot be generated with conventional electrical or mechanical switching techniques [11], [13]. To outline the concept behind ultra-short pulse generation, we will use a simple mathematical model that describes the complex spectrum

$$\underline{E}(\omega) = \underline{A}(\omega) \sum_{k=-\infty}^{\infty} \delta\left(\omega - \omega_{\mathrm{c}} - k\omega_{\mathrm{rt}}\right) \tag{1.3}$$

of the electric field of a Fabry-Pérot Laser (cf. the spectrum in Fig. 1.2), where the complex function $\underline{A}(\omega)$ represents the amplitudes of the different modes and typically features a certain bandwidth around a carrier frequency ω_{c}. The Dirac distribution $\delta(\omega)$ provides an idealization of each mode of the laser, where the line shape has been neglected for simplicity. The infinite sum over the Dirac distributions

$$\sum_{k=-\infty}^{\infty} \delta\left(\omega - k\omega_{\mathrm{rt}}\right) = \mathcal{F}\left\{\frac{1}{\omega_{\mathrm{rt}}} \sum_{k=-\infty}^{\infty} \delta\left(t - kT_{\mathrm{rt}}\right)\right\} \tag{1.4}$$

is sometimes called a Dirac comb and features the interesting property that it is its own Fourier transform [14], which is denoted with $\mathcal{F}\{\cdot\}$. The complex spectrum in Eq. (1.3) can be transformed to the time domain using the inverse Fourier transform $\mathcal{F}^{-1}\{\cdot\}$, which yields

$$\underline{E}(t) = \underline{a}(t) * \frac{1}{\omega_{\mathrm{rt}}} \sum_{k=-\infty}^{\infty} \delta\left(t - kT_{\mathrm{rt}}\right) \exp\left(\mathrm{i}\omega_{\mathrm{c}} t\right) = \frac{1}{\omega_{\mathrm{rt}}} \exp\left(\mathrm{i}\omega_{\mathrm{c}} t\right) \sum_{k=-\infty}^{\infty} \underline{a}\left(t - kT_{\mathrm{rt}}\right). \tag{1.5}$$

Here, we used the modulation theorem to extract the term $\exp(\mathrm{i}\omega_\mathrm{c}t)$, the aforementioned property of the Dirac comb, and the convolution theorem as well as the general properties of the convolution operation $*$ [14]. The resulting time profile features a carrier signal with the frequency ω_c and an envelope function that is periodic with the round trip time T_rt. The form of the latter is defined by the inverse Fourier transform $\underline{a}(t) = \mathcal{F}^{-1}\{\underline{A}(\omega)\}$ of the mode amplitude function $\underline{A}(\omega)$. We note that if $\underline{a}(t)$ constitutes a pulse with a certain temporal width that is smaller than the round trip time, the time profile in Eq. (1.5) describes a train of ultra-short pulses.

Even without exact knowledge of the mode amplitude function $\underline{A}(\omega)$ we can derive two general statements about whether or not pulses are formed. To this end, we write the complex spectrum

$$\underline{A}(\omega) = |A(\omega)| \exp[\mathrm{i}\phi(\omega)] \tag{1.6}$$

in polar notation using the absolute value $|A(\omega)|$ and the (possibly frequency dependent) phase function $\phi(\omega)$. Basing on the similarity theorem, which relates a broad spectrum in frequency domain to a narrow pulse shape in time domain [14], we can deduce that a short temporal width of $|\underline{a}(t)|$ requires a a large spectral width of $|A(\omega)|$. This first statement is a necessary condition, but not sufficient for the generation of ultra-short pulses. To demonstrate this, we expand the phase function

$$\phi(\omega) = \phi_0 + \phi_1\omega + \phi_2\omega^2 + O(\omega^3) \tag{1.7}$$

by using the coefficients ϕ_l and the Landau notation O for the asymptotic upper bound. Then, we note that for $\phi_{l,l\geq 2} = 0$ the resulting time profile can be written as

$$\underline{E}(t) = \frac{1}{\omega_\mathrm{rt}} \exp(\mathrm{i}\omega_\mathrm{c}t + \phi_0) \sum_{k=-\infty}^{\infty} \underline{a}'(t - kT_\mathrm{rt} + \phi_1) \tag{1.8}$$

by virtue of the shifting theorem [14], where $\underline{a}'(t)$ denotes the inverse Fourier transform of $|A(\omega)|$. Although the phase relation between the carrier signal and the envelope function is altered, and the complete signal is shifted in time, the ability to form ultra-short pulses is not lost. However, for a non-zero ϕ_2 (which may occur due to e.g., dispersion) the electric field in Eq. (1.8) has to be convoluted in time domain with the function $\mathcal{F}^{-1}\{\exp(\mathrm{i}\phi_2\omega^2))\}$, which tends to broaden the pulse shape. The higher-order contributions to the phase function $\phi(\omega)$ will have a similar effect and degrade the pulse shape further. As a consequence, ultra-short pulse generation not only requires many modes to be active but also a constant phase relationship between the active modes.

This second requirement leads us to the term mode-locking, which refers to all measures taken to establish such a constant phase relationship. We distinguish between active and passive mode-locking, where the former requires an external oscillator source. This external source oscillates with the frequency ω_rt and thereby synchronizes the laser modes. Passive mode-locking, on the other hand, is implemented exclusively by adapting the resonator cavity so that any phase mismatch is compensated. This is usually achieved by integrating a saturable absorber into the cavity. Additionally, dispersion compensation structures may be required to maintain the phase relation over one round trip [10], [15], [16].

1.3 Quantum cascade lasers

Quantum cascade lasers are special semiconductor lasers whose gain medium does not consist of a single semiconductor material but rather of thin layers of different semiconductor materials. As already mentioned above, the spacing between the lasing energy levels is not defined anymore by the band gap of the semiconductor material, but can be engineered by designing the layers in the gain medium. Since the thickness of the layers is in the (sub)nanometer regime, the operating principle of those lasers is mainly influenced by quantum mechanical effects (as already suggested by their name). One of the most crucial quantum mechanical effects is the discretization of energy due to the confinement by a certain potential, which explains e.g., the formation of atomic orbitals (which are exploited e.g., in gas lasers). Similarly, the

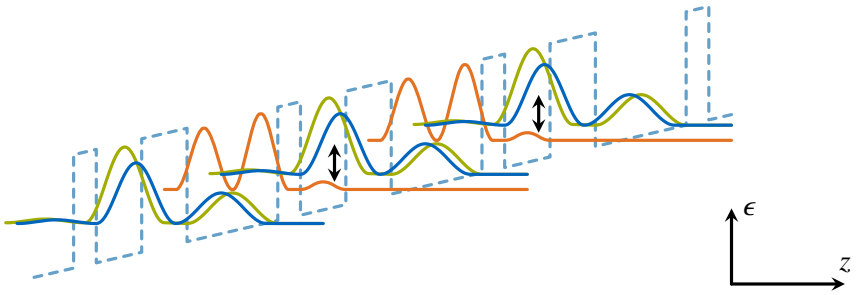


Figure 1.3 Working principle of a quantum cascade laser. The conduction band profile (light blue, dashed) features quantum wells and barriers, which lead to a discretization of the allowed energy values of the electrons. The resulting energy levels are represented by the corresponding wave functions (drawn in solid lines), which depend on the growth direction z. The periodic nature of the conduction band profile allows e.g., the usage the injector level of the next period as ground level. The arrows indicate the lasing transition. Modified from M. Riesch, The QCL Stock Image Project [17] (CC BY 4.0).

layers of a QCL – in conjuction with an external bias voltage, which creates an energy gradient in the gain medium – lead to a discretization of energy that defines the energy levels used for pumping and lasing. It should be pointed out that the layer sequence is periodic and is typically repeated over about 100 periods. Therefore, the energy discretization is periodic as well and resembles a cascade waterfall. This explains the term "quantum cascade" in the name.

Figure 1.3 depicts an exemplary potential in the form of a conduction band profile, which is caused by layers of two different semiconductor materials with different band gap. The resulting conduction band offset creates quantum wells and corresponding barriers that confine the electrons to discrete energy levels. The energy levels are represented by their wave functions (a central concept of quantum mechanics that we will discuss in Chapter 2) in Fig. 1.3, where the offsets of the functions in vertical direction stand for the discrete energy values. For simplicity only three levels are shown to outline the operating principle. Lasing occurs between the upper and lower laser level and the energy difference between those two levels will largely define the frequency of the emitted radiation. The additional injector level should feature a fast non-radiative transition to the upper laser level in order to keep the latter well-populated. In contrast to the pumping scheme in Fig. 1.1, the simplified scheme here does not use an extra depopulation or ground level but rather employs the injector level of the next period to depopulate the lower laser level.

We note that a careful choice of the material system (the combination of different semiconductor materials, e.g., GaAs and AlGaAs) and the doping profile as well as the design of the thickness and sequence of the layers are required to acquire the correct spacing of the energy levels and efficient pumping. These efforts are eventually rewarded with the possibility to emit radiation in the mid-infrared and terahertz regime. The energy spacing that corresponds to those frequency regimes is significantly lower than the band gap of most semiconductor materials. With the approach outlined above, however, suitable energy spacings can be achieved between the so-called intrasubbands (i.e., the energy levels).

In addition to the gain medium including the pumping mechanism, every laser requires a resonator cavity with a feedback mechanism. In the following, a short overview of the QCL subtypes outlines how the resonator cavity is typically realized in QCLs [3], [6]. As an example, Fig. 1.4 shows a typical geometry of a metal-metal wave guide QCL with a Fabry-Pérot cavity. The layer structure of QCLs is fabricated using the molecular beam epitaxy (MBE) or the metalorganic chemical vapor deposition (MOCVD) technique. These processes require a substrate that provides mechanical stability. Usually, QCLs incorporate a wave guide structure that confines the radiation in the transversal directions y and z. This confinement guides the radiation along the propagation direction x. Two wave guide types are used in the majority of QCLs, namely the surface-plasmon (single metal) or the metal-metal (double metal) wave guide. The difference between the two is that the contacts of the metal-metal wave guide completely embed the gain medium of the QCL, whereas the bottom contact of the surface-plasmon wave guide is adjacent to the gain medium, i.e., the gain medium and the substrate are not separated by the bottom contact. For the contacts, a good conductor material is selected (usually copper or gold). The feedback mechanism is often implemented by

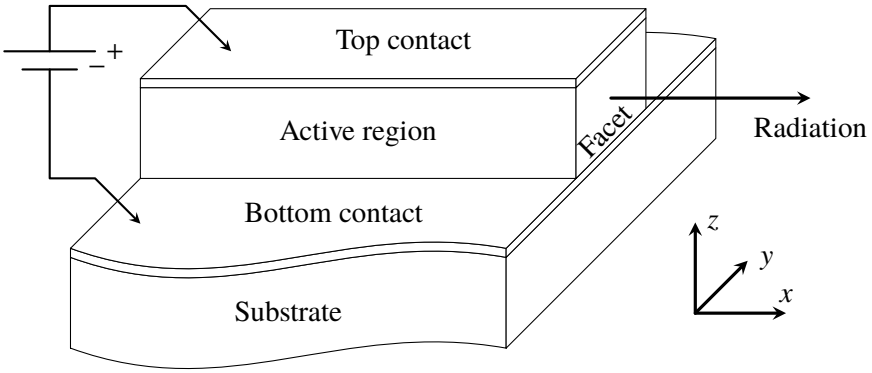


Figure 1.4 Typical geometry of a metal-metal wave guide QCL. The active region contains the gain medium and is embedded between two metal contacts. The constant voltage source delivers the external bias voltage U_B between the contacts. The length in x direction is in the mm regime, the dimensions in y and z direction are usually several tens of µm. Modified from M. Riesch, The QCL Stock Image Project [17] (CC BY 4.0).

cleaving the complete structure in propagation direction x at the desired cavity length. The resulting facets act as semitransparent mirrors and turn the structure into a Fabry-Pérot cavity. For other use cases, QCLs with ring cavities are routinely produced.

Since their first experimental realization, QCLs have undergone an astonishing development [3]. While mid-infrared QCLs are commercially available and routinely used for a longer time, terahertz QCLs turned out to be more complicated to realize. For quite some time the generation of terahertz radiation with QCLs (not counting the difference frequency generation (DFG) approach, where two mid-infrared frequencies were mixed to produce terahertz radiation) has been restricted to cryogenic temperatures [18]. Recent advances in the design of cooling equipment and the gain medium allowed terahertz QCLs to operate at temperatures that are in the reach of thermoelectric Peltier cooling elements. These encouraging developments suggest that terahertz QCLs may be operated without the need for cryogenic equipment in near future, thus enabling various applications such as terahertz spectroscopy [5]. The applications of QCLs often require frequency comb operation [19] or ultra-short pulses [6]. While active mode-locking of QCLs is meanwhile applied routinely, passive mode-locking has not been realized yet. Additionally, even more elaborate operating modes, such as pulse generation with harmonically mode-locked lasers, have become increasingly interesting [20].

1.4 Overview of this work

In the introduction above the basic operating principles of lasers in general and quantum cascade lasers in particular have been outlined. Additionally, we have discussed advanced operating modes of lasers, especially frequency comb operation and ultra-short pulse generation, and have given a short overview of the possible applications of QCLs operating in these modes. Although significant success has been achieved over the last decades, there are still unresolved issues that impede, e.g., the passive mode-locking of QCLs. Seeing that the achievements in experimental work went hand in hand with rigorous theoretical modeling and extensive computer simulations, it seems more than reasonable to use the latter two to extend the understanding of the dynamical processes in the QCLs that is required to exploit these advanced operating modes. Indeed, a variety of studies has been published that investigate the dynamical behavior of QCLs using computer simulations. The research on coherent instabilities [21], active mode-locking [22], frequency comb generation [23], the impact of laser seeding on the QCL operation [24], and colliding pulse mode-locking in multi-section QCL devices [25] are only few of many examples. After reading the related publications it soon becomes apparent that most research groups develop and maintain their own closed-source simulation tools. As we shall see later in this thesis, this requires significant efforts that are carried out multiple times by the individual research groups. In order to reduce the amount of duplicated work, a flexible open-source software framework for the computer-aided simulation of the dynamical behavior

of QCLs is presented in the thesis at hand. This framework is envisaged to provide a solid basis for the upcoming endeavors to the scientific community and to ease the collaboration between the research groups.

In the following, the path to such a flexible framework from the theoretical basis to the verification of the simulation results is outlined. Each step on this path corresponds to a chapter in this thesis and is described briefly at this point.

Chapter 2 provides the necessary theoretical background that is required during the subsequent chapters. Since the main product of any laser is electromagnetic radiation, we begin with an introduction to the theory of electrodynamics. The dynamical processes in the gain medium of a QCL can only be explained with quantum mechanics, which is treated subsequently. The chapter concludes with a concise introduction to the application of quantum mechanics to describe semiconductor materials and semiconductor heterostructures, as these are the essential component of a QCL gain medium.

Chapter 3 bases on the theoretical background and describes a mathematical model of the dynamical behavior of QCLs. In order to provide an efficient approach, static prerequisite simulations that have to be run only once for a certain device are extracted from the model. After this extraction, the generalized Maxwell-Bloch equations are identified as central component of the remaining dynamic model. Additionally, software tools for the prerequisite simulations are reviewed.

Chapter 4 presents an open-source solver for the generalized Maxwell-Bloch equations. This solver is implemented in C++ with a user-friendly Python interface and provides a flexible framework to test different numerical methods and parallelization techniques, and to run productive simulations. Apart from the implementation details and design rationale of this software project, the chapter discusses different test cases that are used to verify the implementation.

Chapter 5 takes a short detour and summarizes the efforts that are required to create a sustainable open-source software project. It reviews the best practices in scientific software engineering and presents an open-source project skeleton for scientific software projects. An instance of this skeleton can be created with a few clicks and reduce the setup time of a new project significantly.

Chapter 6 reviews existing numerical methods for the generalized Maxwell-Bloch equations. Accurate and efficient numerical methods are desired in any simulation, but are required especially in long-term simulations, which are in the scope of this thesis. During the assessment of numerical methods, special care is given to whether or not the methods preserve the physical properties of the involved quantities. The most promising candidates are then compared in terms of single-thread performance.

Chapter 7 presents techniques to run the simulation code on modern parallel architectures. Thereby, the time required to run simulations can be vastly reduced and long-term simulations are facilitated. In this chapter, two implementations of the simulation code are presented, optimized, and compared using multi-threaded performance tests on a multi-core processor.

Chapter 8 discusses the application of the simulation framework to different problems. The first problem is purely theoretical and serves primarily as verification example. However, the validity of certain approximations in the scope of the simulation of ultra-short pulses is discussed with the help of this problem. The two other problems described in this chapter, however, stem from recent experimental work and provide models for self-starting harmonic mode-locking in QCLs and for frequency comb generation in an external-cavity QCL with a graphene saturable absorber as reflector, respectively.

Finally, Chapter 9 gives a conclusion of the thesis at hand and an outlook on possible future work.

2 Theoretical background

Renton: Right. So we all get old and then we can't hack it anymore. Is that it?
Sick Boy: Yeah.
Renton: That's your theory?
Sick Boy: Yeah. Beautifully f&§!ing illustrated.
— *Renton and Sick Boy, Trainspotting*

Let us begin with a concise introduction to the underlying theory, which serves as foundation for the challenges ahead. Since we aim to describe the dynamical behavior of a laser, we need to describe the electromagnetic radiation, which is the main product of any laser, and the gain medium that produces it. As we learned in the previous chapter, the working principle of the quantum cascade laser (QCL) exploits quantum mechanical effects. This fact raises the question whether the QCL dynamics can be described using classical physics. Indeed, the modeling approach presented and used in the subsequent chapters is a semiclassical one. For the electromagnetic field there is no need for a quantized description (i.e., a description in the context of second quantization), as the number of photons (or in other words, the optical power) is sufficiently large. Therefore, a brief summary of classical electrodynamics is given in Section 2.1. In order to describe the gain medium, however, the dynamical behavior of electrons in a semiconductor device has to be modeled, which is not feasible using classical mechanics. Also, the typical dimensions of the QCL layers lead to a confinement of the electrons in growth direction, which can only be explained by the fundamental laws of quantum mechanics. In Section 2.2, those fundamental laws are described briefly. Subsequently, Section 2.3 outlines the application of quantum mechanics to semiconductors and semiconductor heterostructures, respectively.

It should be noted that this introduction is tailored to make the thesis at hand self-contained and cannot replace the standard text books on the topics involved. For more detailed explanations of the matters discussed in the following, the interested reader is referred to the referenced literature.

2.1 Classical electrodynamics

In this branch of theoretical physics, the dynamical behavior of electric charges, currents, and electromagnetic fields is studied (see the text books of Griffiths [26], Jackson [27], or Jin [28] for a thorough introduction). This behavior is governed by the Lorentz force

$$\boldsymbol{F} = q\left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}\right), \tag{2.1}$$

which is the force that a particle with electric charge q experiences when moving with a velocity $\boldsymbol{v}$ in an electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$, and Maxwell's equations

$$\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B}, \tag{2.2a}$$
$$\nabla \times \boldsymbol{B} = \mu_0\left(\boldsymbol{J} + \varepsilon_0 \partial_t \boldsymbol{E}\right), \tag{2.2b}$$
$$\nabla \boldsymbol{E} = \varepsilon_0^{-1} \rho, \tag{2.2c}$$
$$\nabla \boldsymbol{B} = 0, \tag{2.2d}$$

here given in microscopic form. This version considers the permittivity ε_0 and permeability μ_0 of vacuum, respectively, and is particularly useful for modeling on atomic level. Here, the current density $\boldsymbol{J} = \boldsymbol{J}_\mathrm{b} + \boldsymbol{J}_\mathrm{f}$

and charge density $\rho = \rho_\mathrm{b} + \rho_\mathrm{f}$ contain both contributions from bound and freely moving particles. By taking the divergence of Eq. (2.2b), which reads

$$\nabla (\nabla \times \boldsymbol{B}) = 0 = \mu_0 (\nabla \boldsymbol{J} + \varepsilon_0 \partial_t \nabla \boldsymbol{E}), \tag{2.3}$$

and plugging in Eq. (2.2c), we find that there is a constraint

$$\nabla \boldsymbol{J} = -\partial_t \rho \tag{2.4}$$

on the definitions of the current density and the charge density, which is commonly referred to as the continuity equation.

2.1.1 Maxwell's equations in macroscopic form

Depending on the application, however, alternative representations of Maxwell's equations can be advantageous. By defining the auxiliary fields

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}, \tag{2.5a}$$
$$\boldsymbol{H} = \mu_0^{-1} \boldsymbol{B} - \boldsymbol{M}, \tag{2.5b}$$

where $\boldsymbol{P}$ is the polarization density and $\boldsymbol{M}$ is the magnetization field, we can derive a version that is more suitable for macroscopic modeling, e.g., for the treatment of bulk materials. In order to do so, we use

$$\nabla \boldsymbol{D} = \varepsilon_0 \nabla \boldsymbol{E} + \nabla \boldsymbol{P} = \rho_\mathrm{b} + \rho_\mathrm{f} + \nabla \boldsymbol{P}, \tag{2.6}$$
$$\partial_t \boldsymbol{D} = \varepsilon_0 \partial_t \boldsymbol{E} + \partial_t \boldsymbol{P}, \tag{2.7}$$
$$\nabla \times \boldsymbol{H} = \mu_0^{-1} \nabla \times \boldsymbol{B} - \nabla \times \boldsymbol{M} = \boldsymbol{J}_\mathrm{b} + \boldsymbol{J}_\mathrm{f} + \varepsilon_0 \partial_t \boldsymbol{E} - \nabla \times \boldsymbol{M}, \tag{2.8}$$

as well as the definitions $\boldsymbol{J}_\mathrm{b} = \nabla \times \boldsymbol{M} + \partial_t \boldsymbol{P}$ and $\rho_\mathrm{b} = -\nabla \boldsymbol{P}$, to rewrite Eqs. (2.2b) and (2.2c). This yields Maxwell's equations in macroscopic form

$$\nabla \times \boldsymbol{E} = -\partial_t \boldsymbol{B}, \tag{2.9a}$$
$$\nabla \times \boldsymbol{H} = \boldsymbol{J}_\mathrm{f} + \partial_t \boldsymbol{D}, \tag{2.9b}$$
$$\nabla \boldsymbol{D} = \rho_\mathrm{f}, \tag{2.9c}$$
$$\nabla \boldsymbol{B} = 0, \tag{2.9d}$$

where the exact form of the polarization $\boldsymbol{P}$, the magnetization $\boldsymbol{M}$, the current density $\boldsymbol{J}_\mathrm{f}$, and the charge density ρ_f remain to be specified for the respective problem under consideration.

The specifications that define the relationship between the electric field $\boldsymbol{E}$ and the auxiliary field $\boldsymbol{D}$, as well as the relation between their magnetic counterparts $\boldsymbol{B}$ and $\boldsymbol{H}$, are usually referred to as constitutive relations. Generally, the polarization density

$$\boldsymbol{P} = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \sum_i \boldsymbol{\mu}_i \tag{2.10}$$

and the magnetization

$$\boldsymbol{M} = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \sum_i \boldsymbol{m}_i \tag{2.11}$$

collect the individual contributions of all electric dipoles $\boldsymbol{\mu}_i$ and magnetic dipoles $\boldsymbol{m}_i$, respectively, in an infinitesimally small test volume ΔV. For example, the electric dipoles $\boldsymbol{\mu}_i = q_i \boldsymbol{d}_i$ are the product of the charge q_i and the displacement $\boldsymbol{d}_i$ of each dipole, where the latter can be either interpreted as dipole length (i.e., the distance between the positive and the negative charge) or the displacement $\boldsymbol{r}_i - \boldsymbol{r}_0$ of the charge in

relation to a observer point $\boldsymbol{r}_0$. For a continuous distribution $\rho_\mathrm{b} = \rho_\mathrm{b}^+ - \rho_\mathrm{b}^-$ of bound charges within the test volume ΔV, the dipole length can be determined as the distance

$$\boldsymbol{d} = \boldsymbol{d}^+ - \boldsymbol{d}^- = \frac{1}{q}\int_{\Delta V} \rho_\mathrm{b}^+(\boldsymbol{r})\boldsymbol{r}\,\mathrm{d}\boldsymbol{r} - \frac{1}{q}\int_{\Delta V} \rho_\mathrm{b}^-(\boldsymbol{r})\boldsymbol{r}\,\mathrm{d}\boldsymbol{r} = \frac{1}{q}\int_{\Delta V} \rho_\mathrm{b}(\boldsymbol{r})\boldsymbol{r}\,\mathrm{d}\boldsymbol{r} \tag{2.12}$$

between the centers of gravity $\boldsymbol{d}^\pm$ of the positive and negative bound charges $\rho_\mathrm{b}^\pm$, where it is assumed that the total values of positive and negative charges in the test volume are equal to

$$q = \int_{\Delta V} \rho_\mathrm{b}^\pm(\boldsymbol{r})\mathrm{d}\boldsymbol{r}. \tag{2.13}$$

As a consequence, the test volume is charge neutral from a macroscopic point of view. In this case, the polarization can be written as

$$\boldsymbol{P} = \lim_{\Delta V \to 0} \frac{1}{\Delta V}\int_{\Delta V} \rho_\mathrm{b}(\boldsymbol{r})\boldsymbol{r}\,\mathrm{d}\boldsymbol{r} \tag{2.14}$$

For linear and isotropic materials, however, it is easier to use the relations

$$\boldsymbol{D} = \varepsilon_0\boldsymbol{E} + \varepsilon_0\chi_\mathrm{e}\boldsymbol{E} = \varepsilon_0\varepsilon_\mathrm{r}\boldsymbol{E}, \tag{2.15a}$$

$$\boldsymbol{H} = \mu_0^{-1}\boldsymbol{B} - \chi_\mathrm{m}\boldsymbol{H} = \mu_0^{-1}\mu_\mathrm{r}^{-1}\boldsymbol{B}, \tag{2.15b}$$

where χ_e and χ_m are the electric and magnetic susceptibilities, respectively, $\varepsilon_\mathrm{r} = 1 + \chi_\mathrm{e}$ is the relative permittivity, and $\mu_\mathrm{r} = 1 + \chi_\mathrm{m}$ is the relative permeability.

Finally, for the current density $\boldsymbol{J}_\mathrm{f}$, Ohm's law in the form

$$\boldsymbol{J}_\mathrm{f} = \sigma\boldsymbol{E} \tag{2.16}$$

is one possible relation, where σ denotes the electric conductivity.

2.1.2 Maxwell's equations in time-harmonic form

Often, it makes sense to transform Maxwell's equations to frequency domain. This can be achieved by choosing the ansatz $\boldsymbol{C} = \Re\{\underline{\mathbf{C}}^\omega \exp(-\mathrm{i}\omega t)\}$ for the fields $\boldsymbol{C} \in \{\boldsymbol{B}, \boldsymbol{D}, \boldsymbol{E}, \boldsymbol{H}\}$, where ω denotes the angular frequency. At this point it should be noted that the physics convention for time-harmonic quantities is used in this thesis. Plugging the time-harmonic ansatz into Eqs. (2.9) yields

$$\nabla \times \underline{\mathbf{E}}^\omega = \mathrm{i}\omega\underline{\mathbf{B}}^\omega, \tag{2.17a}$$

$$\nabla \times \underline{\mathbf{H}}^\omega = \underline{\mathbf{J}}_\mathrm{f}^\omega - \mathrm{i}\omega\underline{\mathbf{D}}^\omega, \tag{2.17b}$$

$$\nabla\underline{\mathbf{D}}^\omega = \underline{\rho}_\mathrm{f}^\omega, \tag{2.17c}$$

$$\nabla\underline{\mathbf{B}}^\omega = 0, \tag{2.17d}$$

where we assume the implicit transformation of the current density $\boldsymbol{J}_\mathrm{f}$ and the charge density ρ_f to frequency domain.

A more compact form of Eq. (2.17b) can be obtained for the constitutive relations in Eqs. (2.15a) and (2.16). Inserting the relations yields

$$\nabla \times \underline{\mathbf{H}}^\omega = \sigma\underline{\mathbf{E}}^\omega - \mathrm{i}\omega\varepsilon_0\varepsilon_\mathrm{r}\underline{\mathbf{E}}^\omega = -\mathrm{i}\omega\varepsilon_0\underline{\varepsilon}_\mathrm{r}\underline{\mathbf{E}}^\omega, \tag{2.18}$$

where the complex relative permittivity $\underline{\varepsilon}_\mathrm{r} = \varepsilon_\mathrm{r} + \mathrm{i}\sigma/(\omega\varepsilon_0)$ includes both relative permittivity and conductivity.

2.1.3 Maxwell's equations in potential formulation

In the form as described above Maxwell's equations are redundant. This redundancy can be demonstrated and eliminated by introducing the scalar potential φ and the vector potential $\boldsymbol{A}$, and defining the relations

$$\boldsymbol{E} = -\nabla\varphi - \partial_t \boldsymbol{A}, \tag{2.19a}$$

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}. \tag{2.19b}$$

We insert these relations into Eqs. (2.2a) and (2.2d), which yields

$$\nabla \times \boldsymbol{E} = -\nabla \times (\nabla\varphi + \partial_t \boldsymbol{A}) = -\partial_t(\nabla \times \boldsymbol{A}) = -\partial_t \boldsymbol{B}, \tag{2.20}$$

since the curl operation on a gradient field yields zero, and

$$\nabla \boldsymbol{B} = \nabla(\nabla \times \boldsymbol{A}) = 0, \tag{2.21}$$

since the divergence of a rotation field is zero. We can readily see that the equations have been reduced to identities. We proceed in similar fashion with Eq. (2.2b), which results in

$$\nabla \times \boldsymbol{B} = \nabla \times \nabla \times \boldsymbol{A} = \mu_0 (\boldsymbol{J} + \varepsilon_0 \partial_t \boldsymbol{E}) = \mu_0 \left[\boldsymbol{J} - \varepsilon_0 \left(\partial_t \nabla\varphi + \partial_t^2 \boldsymbol{A}\right)\right]. \tag{2.22}$$

After further rearrangements, including using the identity $\nabla \times \nabla \times \boldsymbol{A} = \nabla(\nabla \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$, this equation reads

$$\nabla^2 \boldsymbol{A} - \varepsilon_0 \mu_0 \partial_t^2 \boldsymbol{A} - \nabla (\nabla \boldsymbol{A} + \varepsilon_0 \mu_0 \partial_t \varphi) = -\mu_0 \boldsymbol{J}. \tag{2.23}$$

Lastly, we transform Eq. (2.2c) to

$$\nabla \boldsymbol{E} = -\nabla^2 \varphi - \partial_t \nabla \boldsymbol{A} = \varepsilon_0^{-1} \rho \tag{2.24}$$

and note that we have effectively replaced the vector fields $\boldsymbol{E}$ and $\boldsymbol{B}$ (totaling to six components) with the scalar field φ and the vector field $\boldsymbol{A}$, which consist of only four components and are governed by the Eqs. (2.23) and (2.24).

2.1.4 Gauge invariance of Maxwell's equations

The scalar field φ and the vector field $\boldsymbol{A}$ introduced above are only mathematical utilities rather than physically meaningful quantities. As such, they can be varied as long as the physical results are not affected. This can be demonstrated by introducing a new set of potentials

$$\varphi' = \varphi - \partial_t \xi, \tag{2.25a}$$

$$\boldsymbol{A}' = \boldsymbol{A} + \nabla\xi, \tag{2.25b}$$

where ξ is an arbitrary twice-differentiable function. We note that the electric field

$$\boldsymbol{E} = -\nabla\varphi' - \partial_t \boldsymbol{A}' = -\nabla\varphi + \partial_t \nabla\xi - \partial_t \boldsymbol{A} - \partial_t \nabla\xi = -\nabla\varphi - \partial_t \boldsymbol{A} \tag{2.26}$$

as well as the magnetic field

$$\boldsymbol{B} = \nabla \times \boldsymbol{A}' = \nabla \times \boldsymbol{A} + \nabla \times \nabla\xi = \nabla \times \boldsymbol{A} \tag{2.27}$$

are invariant of the choice of ξ. For example, we can define the function ξ to be the solution of

$$\nabla^2 \xi = -\nabla \boldsymbol{A} \tag{2.28}$$

with $\nabla \boldsymbol{A}' = 0$ as a consequence. Then, we insert the potentials $\varphi = \varphi' + \partial_t \xi$ and $\boldsymbol{A} = \boldsymbol{A}' - \nabla\xi$ into Eqs. (2.23) and (2.24), which yields

$$\nabla^2 \boldsymbol{A}' + \nabla (\nabla \boldsymbol{A}) - \varepsilon_0 \mu_0 \partial_t^2 \boldsymbol{A}' + \varepsilon_0 \mu_0 \partial_t^2 \nabla\xi - \nabla (\nabla \boldsymbol{A}) - \varepsilon_0 \mu_0 \partial_t \nabla (\varphi' + \partial_t \xi) = -\mu_0 \boldsymbol{J} \tag{2.29}$$

and

$$-\nabla^2\varphi' - \partial_t\nabla^2\xi - \partial_t\nabla\boldsymbol{A}' = \varepsilon_0^{-1}\rho, \tag{2.30}$$

respectively. After removing the terms that cancel each other, the equations read

$$\nabla^2\boldsymbol{A}' - \varepsilon_0\mu_0\partial_t^2\boldsymbol{A}' - \varepsilon_0\mu_0\partial_t\nabla\varphi' = -\mu_0\boldsymbol{J}, \tag{2.31a}$$

$$-\nabla^2\varphi' = \varepsilon_0^{-1}\rho, \tag{2.31b}$$

and are referred to as the potential formulation of Maxwell's equations in Coulomb gauge.

2.2 Fundamentals of quantum mechanics

Similar to other scientific theories, quantum mechanics bases on a small set of postulates, which we introduce in the following. In this section, we follow the descriptions given by Tang [29] in his very well-written introduction to quantum mechanics, and by Nielsen and Chuang [30] in their comprehensive book on quantum computation and quantum information.

2.2.1 Quantum mechanical states and operators

In classical mechanics, the state of a system at any given time can be described completely by the position $\boldsymbol{r}$ and the momentum $\boldsymbol{p}$. This description changes substantially in the scope of quantum mechanics, as we shall see.

Postulate 1 The state of a dynamic system is completely described by a state vector $|\Psi\rangle$, which is a unit vector of the state space that is associated with the system. The state space is a complex vector space with an inner product $\langle\cdot|\cdot\rangle$, i.e., a Hilbert space.

In order to support the intentionally abstract formulation of the postulate, we use the Dirac notation for the state vectors and the inner product. This notation is independent of their exact mathematical form, which can be specified at a later stage, most commonly when discussing a particular problem. For example, if the problem depends on three spatial dimensions, one can write the state vector $|\Psi\rangle$ as complex function $\Psi(\boldsymbol{r})$, where $\boldsymbol{r}$ is the position in space. In this case, the corresponding inner product assumes the form

$$\langle\Psi|\Psi\rangle = \int \Psi^*(\boldsymbol{r})\Psi(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \int |\Psi(\boldsymbol{r})|^2\,\mathrm{d}\boldsymbol{r} = 1 \tag{2.32}$$

and must be equal to unity as postulated above. Here, the term $|\Psi(\boldsymbol{r})|^2$ can be interpreted as the probability distribution function of finding a particle in the state $|\Psi\rangle$ at a certain position $\boldsymbol{r}$, and the normalization condition in Eq. (2.32) guarantees that the particle can be found somewhere in space.

Although the state vector $|\Psi\rangle$ contains the information that describes a certain state completely, it is not exactly meaningful in a physical sense. Therefore, we need a further entity that extracts physically meaningful quantities from the state vector.

Postulate 2 Every physically observable property Q of a dynamic system is represented by a linear operator $\hat{Q}$ that acts on the state space. The measurement of Q changes the state of the system to another state $|\Psi'\rangle = \hat{Q}\,|\Psi\rangle$. In contrast to classical mechanics, the result of the measurement has only statistical meaning and is given as expectation value $\langle\hat{Q}\rangle = \langle\Psi|\hat{Q}|\Psi\rangle$.

The form of the operators depends on the chosen representation. In position space, we can use the definitions $\hat{\boldsymbol{r}} = \boldsymbol{r}$ and $\hat{\boldsymbol{p}} = -\mathrm{i}\hbar\nabla$ for the position and the momentum operator, respectively, where $\hbar$ is the reduced Planck constant. An equivalent formulation can be formed in momentum space. Independent of the chosen space, any other operator representing a physical quantity can be derived in the same way the quantity would be derived from position and momentum in classical terms.

The introduction of quantum mechanical states and operators reflects a crucial feature of systems with small dimensions, namely the impact of the measurement on the system state. As a consequence, the order

of two different measurements may make result in different outcome. For example, the measurement of position and momentum in different order

$$\hat{\boldsymbol{r}}\,\hat{\boldsymbol{p}}\,|\Psi\rangle = -\mathrm{i}\hbar\hat{\boldsymbol{r}}\nabla\,|\Psi\rangle \neq \hat{\boldsymbol{p}}\,\hat{\boldsymbol{r}}\,|\Psi\rangle = -\mathrm{i}\hbar\nabla\hat{\boldsymbol{r}}\,|\Psi\rangle = -\mathrm{i}\hbar\,|\Psi\rangle - \mathrm{i}\hbar\hat{\boldsymbol{r}}\nabla\,|\Psi\rangle \tag{2.33}$$

leads to the commutation relation

$$\hat{\boldsymbol{r}}\,\hat{\boldsymbol{p}} - \hat{\boldsymbol{p}}\,\hat{\boldsymbol{r}} = [\hat{\boldsymbol{r}}, \hat{\boldsymbol{p}}] = \mathrm{i}\hbar, \tag{2.34}$$

which is a manifestation of Heisenberg's uncertainty principle. Here, $[\cdot,\cdot]$ denotes the commutator. In the case that the commutator of two operators equals zero, the order of the measurements does not make a difference and the corresponding quantities are not subject to the uncertainty principle.

We note that if the state vector $|\Psi\rangle$ is an eigenvector of the operator $\hat{Q}$, then $\hat{Q}\,|\Psi\rangle = \lambda\,|\Psi\rangle$ holds, where λ is the corresponding eigenvalue. In this special case, the measurement does not change the state of the system and the expectation value

$$\langle\hat{Q}\rangle = \langle\Psi|\hat{Q}|\Psi\rangle = \langle\Psi|\lambda|\Psi\rangle = \lambda\,\langle\Psi|\Psi\rangle = \lambda \tag{2.35}$$

equals the eigenvalue. As the expectation value is always related to a physical quantity, it must be real-valued. Consequently, all eigenvalues of the operators under consideration must be real-valued, which is the case for Hermitian operators $\hat{Q} = \hat{Q}^\dagger$.

Frequently, the quantum mechanical states and operators are projected onto the eigenvectors $|\phi_i\rangle$ of the operators. In the resulting representation, the state vector

$$|\Psi\rangle = \sum_i \langle\phi_i|\Psi\rangle\,|\phi_i\rangle = \sum_i c_i\,|\phi_i\rangle \tag{2.36}$$

can be written as linear combination of the basis functions and it is sufficient to consider the coefficients c_i in a vector. Similarly, the state vector $|\Psi'\rangle = \hat{Q}\,|\Psi\rangle$ can be expanded as

$$|\Psi'\rangle = \sum_i \langle\phi_i|\hat{Q}|\Psi\rangle\,|\phi_i\rangle = \sum_{i,j} \langle\phi_i|\hat{Q}|\phi_j\rangle\,\langle\phi_j|\Psi\rangle\,|\phi_i\rangle = \sum_i c_i'\,|\phi_i\rangle\,, \tag{2.37}$$

where the coefficients $c_i' = \sum_j q_{ij}c_j$ depend on the matrix elements $q_{ij} = \langle\phi_i|\hat{Q}|\phi_j\rangle$ and the coefficients c_i. As a result, we can write the state vector $|\Psi\rangle$ as actual vector and the operator $\hat{Q}$ as matrix. In the following, we shall use the terms operator and matrix interchangeably.

2.2.2 Time evolution of quantum mechanical states

Now that we have defined quantum mechanical states and operators, we can take a look on how they evolve with respect to time. It should be noted, however, that there is more than one way to look at this time evolution. Unless noted otherwise, the Schrödinger picture is used in this thesis. Here, the state vectors evolve in time but the operators are not explicitly time dependent. In the Heisenberg picture the opposite is true. Finally, both state vectors and operators evolve in time in the interaction picture.

Postulate 3 All state vectors $|\Psi\rangle$ satisfy the time-dependent Schrödinger equation

$$\mathrm{i}\hbar\partial_t\,|\Psi\rangle = \hat{H}\,|\Psi\rangle\,, \tag{2.38}$$

where the Hamiltonian operator $\hat{H} = \hat{T} + \hat{V}$ is the sum of kinetic energy $\hat{T}$ and potential energy $\hat{V}$ and therefore represents the total energy of the system. Here, $\hbar$ denotes the reduced Planck constant.

The Hamiltonian function $\mathcal{H}$ is the centerpiece of the eponymous formulation of mechanics, which has been developed to provide a generalized approach to classical mechanics. However, it can be transferred directly to quantum mechanics by replacing the physical quantities with their corresponding quantum mechanical operators. In the scope of this formulation, the change in position and momentum can be determined with Hamilton's equations

$$\mathrm{d}_t r_i = \partial_{p_i}\mathcal{H}, \tag{2.39a}$$

$$\mathrm{d}_t p_i = -\partial_{r_i}\mathcal{H}, \tag{2.39b}$$

where $r_i \in \{x, y, z\}$ and $p_i \in \{p_x, p_y, p_z\}$ denote the components of the position and momentum, respectively, and $\mathrm{d}_t r_i$, $\mathrm{d}_t p_i$ represent their total derivatives with respect to time.

Using those equations, we can easily verify that the Hamiltonian operator

$$\hat{H}_0 = \frac{\hat{\boldsymbol{p}}^2}{2m} + V = -\frac{\hbar^2 \nabla^2}{2m} + V \tag{2.40}$$

describes the behavior of a system with mass m in a conservative potential V, since the expressions for velocity and force

$$\mathrm{d}_t r_i = \frac{p_i}{m} \rightarrow \mathrm{d}_t \boldsymbol{r} = \boldsymbol{v} = \frac{\boldsymbol{p}}{m}, \tag{2.41a}$$

$$\mathrm{d}_t p_i = -\partial_{r_i} V \rightarrow \mathrm{d}_t \boldsymbol{p} = m\mathrm{d}_t^2 \boldsymbol{r} = \boldsymbol{F} = -\nabla V \tag{2.41b}$$

yield the well-known equations $\boldsymbol{p} = m\boldsymbol{v}$ and $\boldsymbol{F} = -\nabla V$. Note that quantum mechanical operators and the corresponding physical quantities are used interchangeably and position and momentum are considered independent of each other.

The solution of the Schrödinger equation (2.38) can be generally given in the form

$$|\Psi(t)\rangle = \hat{U}(t, t_0)\,|\Psi(t_0)\rangle \tag{2.42}$$

using the unitary time evolution operator

$$\hat{U}(t, t_0) = \mathcal{T}_{\leftarrow} \exp\left[-\mathrm{i}\hbar^{-1} \int_{t_0}^{t} \hat{H} \mathrm{d}\tau\right], \tag{2.43}$$

where the time-ordering operator $\mathcal{T}_{\leftarrow}$ ensures the correct order of time-dependent operators in products. In the case of a time-independent Hamiltonian $\hat{H} = \hat{H}_0$, the evolution operator can be simplified significantly to yield

$$\hat{U}_0(t, t_0) = \exp\left[-\mathrm{i}\hbar^{-1} \hat{H}_0 (t - t_0)\right]. \tag{2.44}$$

We can readily see that $\hat{U}\hat{U}^\dagger = \hat{I}$ holds and the time evolution operator is indeed unitary.

For a time-independent potential V (and, consequently, a time-independent Hamiltonian $\hat{H}_0$), the complexity of the Schrödinger equation can be reduced by the separation ansatz $|\Psi\rangle = \Psi(\boldsymbol{r}, t) = R(\boldsymbol{r})T(t)$. Plugging this ansatz and the Hamiltonian $\hat{H}_0$ into Eq. (2.38) yields

$$R(\boldsymbol{r})\mathrm{i}\hbar\partial_t T(t) = T(t)\hat{H}_0 R(\boldsymbol{r}) \rightarrow \mathrm{i}\hbar \frac{\partial_t T(t)}{T(t)} = \frac{\hat{H}_0 R(\boldsymbol{r})}{R(\boldsymbol{r})} = \epsilon, \tag{2.45}$$

where ϵ denotes the separation constant. After rewriting the equation for the spatially dependent component as

$$\hat{H}_0 R(\boldsymbol{r}) = \epsilon R(\boldsymbol{r}), \tag{2.46}$$

which is commonly referred to as the time-independent or stationary Schrödinger equation, it turns out that the separation constant is an eigenvalue of the Hamiltonian operator. Since the Hamiltonian is related to the total energy of the system, those eigenvalues are called the eigenenergies of the system. In this simple case, the time evolution of the corresponding eigenstates can be derived by rewriting Eq. (2.45) as

$$\mathrm{i}\hbar\partial_t T(t) = \epsilon T(t) \rightarrow T(t) = \exp\left[-\mathrm{i}\hbar^{-1}\epsilon(t - t_0)\right] T(t_0). \tag{2.47}$$

2.2.3 Hamiltonian of a system in a classical electromagnetic field

Let us now consider a quantum mechanical system in a classical electromagnetic field, which is a quite common problem in quantum mechanics. In this case, the Hamiltonian can be taken from literature [31] and reads

$$\hat{H} = \frac{(\hat{\boldsymbol{p}} - q\boldsymbol{A})^2}{2m} + q\varphi + V \tag{2.48}$$

where q is the electric charge of the system, and the potentials $\boldsymbol{A}$ and φ introduced in Eqs. (2.19) describe the electromagnetic field. Then, Hamilton's equations yield

$$\mathrm{d}_t r_i = \frac{p_i - qA_i}{m} \rightarrow \mathrm{d}_t \boldsymbol{r} = \boldsymbol{v} = \frac{\boldsymbol{p} - q\boldsymbol{A}}{m}, \tag{2.49a}$$

$$\mathrm{d}_t p_i = \frac{\boldsymbol{p} - q\boldsymbol{A}}{m} \left(q\partial_{r_i} \boldsymbol{A}\right) - \partial_{r_i} \left(q\varphi + V\right). \tag{2.49b}$$

By plugging the total time derivative of Eq. (2.49a) into Eq. (2.49b), we can determine the acting force

$$F_i = m\mathrm{d}_t^2 r_i = \mathrm{d}_t p_i - q\mathrm{d}_t A_i = \boldsymbol{v} \left(q\partial_{r_i} \boldsymbol{A}\right) - \partial_{r_i} \left(q\varphi + V\right) - q\mathrm{d}_t A_i \tag{2.50}$$

elementwise. Using the relation $\mathrm{d}_t A_i = \partial_t A_i + \boldsymbol{v}\nabla A_i$ for the total time derivative of the vector field $\boldsymbol{A}$, we can simplify the expression

$$\boldsymbol{v}\partial_{r_i} \boldsymbol{A} - \mathrm{d}_t A_i = \boldsymbol{v}\partial_{r_i} \boldsymbol{A} - \partial_t A_i - \boldsymbol{v}\nabla A_i = -\partial_t A_i + (\boldsymbol{v} \times \nabla \times \boldsymbol{A})_i \tag{2.51}$$

and write the force

$$\boldsymbol{F} = q\left(-\nabla\varphi - \partial_t \boldsymbol{A} + \boldsymbol{v} \times \nabla \times \boldsymbol{A}\right) - \nabla V \tag{2.52}$$

as vector. After comparison with Eqs. (2.1) and (2.19), we can readily see that the acting force contains the Lorentz force in potential formulation.

We note that the canonical momentum $\boldsymbol{p'} = m\boldsymbol{v} + q\boldsymbol{A}$ in Eq. (2.48) deviates from the mechanical momentum $m\boldsymbol{v}$, which renders the physical interpretation of results complicated [31], [32]. As a remedy, the Hamiltonian can be transformed to a version where the canonical and mechanical momentum coincide, such as in the Hamiltonian $\hat{H}_0$ in Eq. (2.40). This procedure requires a certain assumption, which we shall discuss in Section 3.1.3. For the time being, the transformation itself is described. This can be accomplished either in the scope the Lagrangian formalism of mechanics [33] or basing on the Hamiltonian formalism, where the procedure is referred to as canonical transformation. In the following, let us consider an auxiliary function $\xi(\boldsymbol{r}, t)$, the coordinates $\boldsymbol{r} = \boldsymbol{r'}$ and $\boldsymbol{p} = \boldsymbol{p'} + \nabla\xi$, as well as the Hamiltonian

$$\hat{H}(\boldsymbol{r}, \boldsymbol{p}, t) = \hat{H}'(\boldsymbol{r'}, \boldsymbol{p'}, t) - \partial_t \xi. \tag{2.53}$$

Naturally, such a transformation is only valid if Hamilton's equations (cf. Eqs. (2.39) for the original form) are not altered for the newly introduced Hamiltonian and coordinates. Since the auxiliary function does not depend on the momentum, we can use $\partial p_i'/\partial p_i = 1$ and write

$$\mathrm{d}_t r_i = \partial_{p_i} \hat{H} = \partial_{p_i'} \hat{H} = \partial_{p_i'} \hat{H}' - \partial_{p_i'} \partial_t \xi = \partial_{p_i'} \hat{H}' = \mathrm{d}_t r_i' \tag{2.54}$$

for the first equation. We proceed similarly with the second equation, which yields

$$\mathrm{d}_t p_i = \mathrm{d}_t p_i' + \mathrm{d}_t \partial_{r_i} \xi = -\partial_{r_i} \hat{H}(\boldsymbol{r}, \boldsymbol{p}, t) = -\partial_{r_i} \hat{H}(\boldsymbol{r}, \boldsymbol{p'}, t) - \partial_{p_i'} \hat{H}(\boldsymbol{r}, \boldsymbol{p'}, t) \partial_{r_i} p_i', \tag{2.55}$$

and note that the case is more complicated due to the fact that the new momentum coordinate depends on the spatial coordinate. We plug in the definition of the new Hamiltonian into the expression, which leads to

$$\mathrm{d}_t p_i' + \mathrm{d}_t \partial_{r_i} \xi = -\partial_{r_i} \hat{H}' + \partial_{r_i} \partial_t \xi - \partial_{p_i'} \hat{H}' \partial_{r_i} p_i'. \tag{2.56}$$

By using the identities $\partial_{r_i} p_i' = -\partial_{r_i}^2 \xi$ and $\partial_{p_i'} \hat{H}' = \mathrm{d}_t r_i' = \mathrm{d}_t r_i$, we find that the total derivative

$$\mathrm{d}_t \partial_{r_i} \xi = \partial_t \partial_{r_i} \xi + (\mathrm{d}_t r_i)\, \partial_{r_i}^2 \xi \tag{2.57}$$

can be subtracted from both sides of Eq. (2.56), which confirms that also the second of Hamilton's equations

$$\mathrm{d}_t p_i' = -\partial_{r_i} \hat{H}' = -\partial_{r_i'} \hat{H}'. \tag{2.58}$$

is invariant to the transformation described above.

2.2.4 Density operator and the Liouville-von Neumann equation

With the help of the three postulates above we can already describe the physical state of a quantum mechanical system with the state vector $|\Psi\rangle$ and its time evolution by solving the Schrödinger equation. At this point, we can readily treat single-particle problems, such as the behavior of an electron in a potential well. For a system that contains more than one particle, however, the complexity increases. Considering the typical charge carrier densities in semiconductor materials, it becomes apparent that it is hardly feasible to know the exact state of each particle. Fortunately, this exact knowledge is not required in most applications, and instead it is sufficient to know the states $|\Psi_k\rangle$ which the particles of the system can be in, and to determine a statistical probability p_k that indicates to what degree those states are occupied by the particles. Then, we can generalize the expression for the expectation value of an operator $\hat{Q}$ to the weighted average

$$\langle \hat{Q} \rangle_{3\mathrm{D}} = n_{3\mathrm{D}} \langle \hat{Q} \rangle = n_{3\mathrm{D}} \sum_k p_k \langle \Psi_k | \hat{Q} | \Psi_k \rangle \tag{2.59}$$

over all particles and all possible states. Here, $n_{3\mathrm{D}}$ is the particle density in the system. As in Eq. (2.36), we can expand the states $|\Psi_k\rangle$ as linear combination of a basis ϕ_i and rewrite the expectation value as

$$\langle \hat{Q} \rangle = \sum_k p_k \sum_{i,j} \langle \Psi_k | \phi_i \rangle \langle \phi_i | \hat{Q} | \phi_j \rangle \langle \phi_j | \Psi_k \rangle = \sum_{i,j} \langle \phi_i | \hat{Q} | \phi_j \rangle \sum_k p_k \langle \phi_j | \Psi_k \rangle \langle \Psi_k | \phi_i \rangle . \tag{2.60}$$

Now, we introduce the density operator

$$\hat{\rho} = \sum_k p_k |\Psi_k\rangle \langle \Psi_k| \tag{2.61}$$

and find that the expression above can be simplified to

$$\langle \hat{Q} \rangle = \sum_{i,j} \langle \phi_i | \hat{Q} | \phi_j \rangle \langle \phi_j | \hat{\rho} | \phi_i \rangle = \mathrm{Tr}\left\{ \hat{Q} \hat{\rho} \right\}, \tag{2.62}$$

where $\mathrm{Tr}\{\cdot\}$ denotes the trace of a matrix. We note that the density operator effectively replaces several state vectors and allows the compact treatment of systems that cannot be described by a single state.

At this point, a discussion of the probabilities p_k is in order. From probability theory, we expect that $p_k \in [0, 1]$ and the sum over all probability values $\sum_k p_k = 1$. This has important consequences for the density operator $\hat{\rho}$. In the basis of $|\Psi_k\rangle$, $\hat{\rho}$ is a diagonal matrix with the main diagonal entries p_k. This means that the probabilities p_k are the eigenvalues of the density matrix. We can readily see from Eq. (2.61) that the density matrix is Hermitian, which means that the eigenvalues p_k are real. Furthermore, as the p_k are non-negative, we have to demand that the density matrix is positive semidefinite. Finally, the trace of the density matrix

$$\mathrm{Tr}\{\hat{\rho}\} = \sum_k p_k = 1 \tag{2.63}$$

is the sum of the eigenvalues and must sum up to unity.

While we have already derived the expectation value from the density operator, it remains to be shown how the density operator evolves in time. In order to do so, we plug in the Schrödinger equation (2.38) into the expression for the time derivative

$$\partial_t \hat{\rho} = \sum_k p_k \partial_t |\Psi_k\rangle \langle \Psi_k| + p_k |\Psi_k\rangle \partial_t \langle \Psi_k| = -\mathrm{i}\hbar^{-1} \left(\sum_k p_k \hat{H} |\Psi_k\rangle \langle \Psi_k| - \sum_k p_k |\Psi_k\rangle \langle \Psi_k| \hat{H}^\dagger \right) \tag{2.64}$$

which yields the (Liouville-)von Neumann equation

$$\partial_t \hat{\rho} = -\mathrm{i}\hbar^{-1} \left(\hat{H}\hat{\rho} - \hat{\rho}\hat{H} \right) = -\mathrm{i}\hbar^{-1} \left[\hat{H}, \hat{\rho} \right] = \mathcal{L}(\hat{\rho}), \tag{2.65}$$

where the Liouvillian superoperator $\mathcal{L}$ can be used for brevity. In this derivation, we have assumed that the probabilities p_k do not vary in time. As can be seen from the general solution

$$\hat{\rho}(t) = \hat{U}(t, t_0)\, \hat{\rho}(t_0)\, \hat{U}^\dagger(t, t_0) = \mathcal{U}(t, t_0, \hat{\rho}(t_0)) \tag{2.66}$$

of the Liouville-von Neumann equation, the density operator evolves unitarily in time in this case. Here, the unitary time evolution operator from Eq. (2.43) has been used and the solution superoperator $\mathcal{U}$ has been introduced as shortcut. If, on the other hand, we allow for time-dependent probabilities, an extra term arises in Eq. (2.65). This case shall be discussed Section 2.2.6.

2.2.5 Composite systems and the reduced density operator

Up to this point, we implicitly assumed that we are interested in all degrees of freedom of the complete quantum mechanical system. In practice, this might not be the case and a more efficient approach would be beneficial. A common example of such an approach is to divide the complete system in a system of interest S and an environment E. Then, a reduced density operator can be derived that describes the system of interest alone, thereby providing a more compact description. In order to do so, we have to introduce the notion of composite systems in a further postulate.

Postulate 4 The state space of a composite system is the tensor product of the state spaces of the component systems. Furthermore, if the component systems are prepared in the states $|\Psi_i\rangle$, then the joint state of the composite system reads $|\Psi\rangle = |\Psi_1\rangle \otimes |\Psi_2\rangle \otimes \cdots \otimes |\Psi_N\rangle$, where $\otimes$ denotes the tensor product.

As a consequence, if the state spaces of the system under consideration and the environment are spanned by the basis vectors $|\phi_{i,\mathrm{S}}\rangle$ and $|\phi_{j,\mathrm{E}}\rangle$, respectively, we can use the vectors $|\phi_{i,\mathrm{S}}\rangle \otimes |\phi_{j,\mathrm{E}}\rangle$ as basis for the state space of the composite system. Then, any state of the composite system can be written as linear superposition

$$|\Psi\rangle = \sum_{i,j} c_{ij} \, |\phi_{i,\mathrm{S}}\rangle \otimes |\phi_{j,\mathrm{E}}\rangle \tag{2.67}$$

with the coefficients c. At this point, a very interesting implication of the postulate becomes apparent. We note that the joint state can be determined as tensor product of the component states, but the opposite is not necessarily true. In the case in which a decomposition is possible, the joint state is referred to as a separable state. Otherwise, the component systems are said to be entangled, which is a physical phenomena that can only be explained by quantum mechanics.

Similarly, we can write the operators on the composite state space as

$$\hat{Q} = \sum_{i,j,k,l} c'_{ijkl} \left(|\phi_{i,\mathrm{S}}\rangle \langle \phi_{j,\mathrm{S}}| \right) \otimes \left(|\phi_{k,\mathrm{E}}\rangle \langle \phi_{l,\mathrm{E}}| \right) = \sum_{i,j} c''_{ij} \hat{s}_i \otimes \hat{e}_j \tag{2.68}$$

with the coefficients c' and c''. The introduced operators $\hat{s}_i$ and $\hat{e}_j$ constitute a basis of all operators in the reduced system S and the environment E, respectively. Due to the similarity with Eq. (2.67) we deduce that it is possible to compose an operator $\hat{Q} = \hat{Q}_\mathrm{S} \otimes \hat{Q}_\mathrm{E}$ from the component operators, but it is not guaranteed that an operator of the composite system can be decomposed as tensor product of the component operators.

Let us now consider the reduced density operator $\hat{\rho}_\mathrm{S}$ and another operator $\hat{Q}_\mathrm{S}$ that acts on the system under consideration. We know from Eq. (2.62) that the expectation value of the operator $\hat{Q}_\mathrm{S}$ is

$$\langle \hat{Q}_\mathrm{S} \rangle = \mathrm{Tr} \left\{ \hat{Q}_\mathrm{S} \hat{\rho}_\mathrm{S} \right\} . \tag{2.69}$$

With the help of the identity operator $\hat{I}_\mathrm{E}$ of the state space of the environment, we can extend the operator $\hat{Q}_\mathrm{S}$ to an operator $\hat{Q} = \hat{Q}_\mathrm{S} \otimes \hat{I}_\mathrm{E}$ that acts on the composite system. Now we demand that this extended operator must yield the same expectation value, as there is no action within the environment and no interaction between the environment and the reduced system. Then, the equation

$$\langle \hat{Q}_\mathrm{S} \rangle = \mathrm{Tr} \left\{ \hat{Q}_\mathrm{S} \hat{\rho}_\mathrm{S} \right\} \stackrel{!}{=} \mathrm{Tr} \left\{ \hat{Q} \hat{\rho} \right\} \tag{2.70}$$

must hold. In order to verify this equivalence, we need to introduce the partial trace operation

$$\mathrm{Tr}_\mathrm{E} \left\{ \hat{Q}_\mathrm{S} \otimes \hat{Q}_\mathrm{E} \right\} = \hat{Q}_\mathrm{S} \, \mathrm{Tr} \left\{ \hat{Q}_\mathrm{E} \right\} \tag{2.71}$$

with the property $\mathrm{Tr}\{\cdot\} = \mathrm{Tr}_\mathrm{S}\{\mathrm{Tr}_\mathrm{E}\{\cdot\}\}$. With this property we can rewrite the condition as

$$\mathrm{Tr} \left\{ \hat{Q}_\mathrm{S} \hat{\rho}_\mathrm{S} \right\} \stackrel{!}{=} \mathrm{Tr}_\mathrm{S} \left\{ \mathrm{Tr}_\mathrm{E} \left\{ \hat{Q} \hat{\rho} \right\} \right\} \tag{2.72}$$

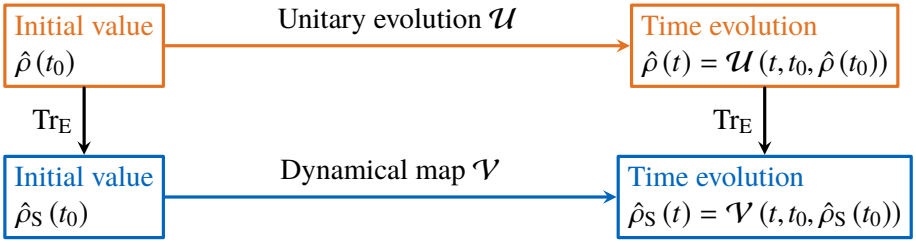


Figure 2.1 Comparison of the initial value and the time evolution of a complete system (closed, in orange) and the system of interest (open, in blue), respectively. Redrawn from H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, 2002 [34].

and note that on the left hand side the trace operation Tr and partial trace operation $\mathrm{Tr_S}$ are equivalent. Hence, we aim to verify the equivalence of the arguments

$$\hat{Q}_\mathrm{S}\hat{\rho}_\mathrm{S} \stackrel{!}{=} \mathrm{Tr_E}\left\{\hat{Q}\hat{\rho}\right\} = \mathrm{Tr_E}\left\{\left(\hat{Q}_\mathrm{S}\otimes\hat{I}_\mathrm{E}\right)\left(\sum_{i,j} c''_{ij}\hat{s}_i\otimes\hat{e}_j\right)\right\} = \mathrm{Tr_E}\left\{\sum_{i,j} c''_{ij}\hat{Q}_\mathrm{S}\hat{s}_i\otimes\hat{e}_j\right\} \tag{2.73}$$

by expanding the density operator $\hat{\rho}$ as described in Eq. (2.68) and using the property $(\hat{A}\otimes\hat{B})(\hat{C}\otimes\hat{D}) = \hat{A}\hat{C}\otimes\hat{B}\hat{D}$ of the tensor product. We can rearrange the resulting expression

$$\mathrm{Tr_E}\left\{\sum_{i,j} c''_{ij}\hat{Q}_\mathrm{S}\hat{s}_i\otimes\hat{e}_j\right\} = \sum_{i,j} c''_{ij}\,\mathrm{Tr_E}\left\{\hat{Q}_\mathrm{S}\hat{s}_i\otimes\hat{e}_j\right\} = \sum_{i,j} c''_{ij}\hat{Q}_\mathrm{S}\hat{s}_i\,\mathrm{Tr}\left\{\hat{e}_j\right\} = \sum_{i,j} c''_{ij}\hat{Q}_\mathrm{S}\,\mathrm{Tr_E}\left\{\hat{s}_i\otimes\hat{e}_j\right\} \tag{2.74}$$

further by using the definition of the partial trace in Eq. (2.71) and find that the condition

$$\hat{Q}_\mathrm{S}\hat{\rho}_\mathrm{S} \stackrel{!}{=} \mathrm{Tr_E}\left\{\hat{Q}\hat{\rho}\right\} = \hat{Q}_\mathrm{S}\,\mathrm{Tr_E}\left\{\hat{\rho}\right\} \tag{2.75}$$

holds if we define the reduced density matrix $\hat{\rho}_\mathrm{S}$ as

$$\hat{\rho}_\mathrm{S} = \mathrm{Tr_E}\left\{\hat{\rho}\right\}. \tag{2.76}$$

2.2.6 Quantum master equations and their solutions

In the section before we have defined the reduced density matrix, which represents the system of interest. The next goal is to describe its time evolution with a so-called quantum master equation, which subsequently replaces the Liouville-von Neumann equation. By applying the time derivative on Eq. (2.76), we can derive

$$\partial_t\hat{\rho}_\mathrm{S} = -\mathrm{i}\hbar^{-1}\,\mathrm{Tr_E}\left\{\left[\hat{H},\hat{\rho}\right]\right\} \tag{2.77}$$

as a starting point for such a master equation. Also, we can construct the Hamiltonian

$$\hat{H} = \hat{H}_\mathrm{S}\otimes\hat{I}_\mathrm{E} + \hat{I}_\mathrm{S}\otimes\hat{H}_\mathrm{E} + \hat{H}_\mathrm{I} \tag{2.78}$$

so that the reduced system alone would be governed by the Hamiltonian $\hat{H}_\mathrm{S}$ (and, analogously, the environment alone by the Hamiltonian $\hat{H}_\mathrm{E}$). The reduced system and the environment are only coupled after the introduction of the interaction Hamiltonian $\hat{H}_\mathrm{I}$. However, deriving a master equation that contains only the reduced density matrix is still far from trivial. In their excellent book about the theory of open quantum systems, Breuer and Petruccione [34] outline two paths to this goal, which will be discussed in the following.

Before, however, it makes sense to discuss the mathematical properties of the solution as a prerequisite for both approaches. After all, the ultimate goal is not only to derive the master equation, but also to solve it. Figure 2.1 compares the initial value and the time evolution of the complete system and the reduced system of interest, respectively. The vertical arrows denote the partial trace operation over the environment, which transforms the initial value of the complete density matrix into that of the reduced density matrix according

to Eq. (2.76). On the right side, the time evolution is transformed by applying the partial trace operation on Eq. (2.66), which yields

$$\hat{\rho}_{\mathrm{S}}(t) = \mathrm{Tr}_{\mathrm{E}}\left\{\hat{U}(t,t_0)\,\hat{\rho}(t_0)\,\hat{U}^{\dagger}(t,t_0)\right\} = \mathrm{Tr}_{\mathrm{E}}\left\{\mathcal{U}(t,t_0,\hat{\rho}(t_0))\right\} = \mathcal{V}(t,t_0,\hat{\rho}_{\mathrm{S}}(t_0)), \tag{2.79}$$

where the dynamical maps $\mathcal{U}$ and $\mathcal{V}$ have been introduced to represent the time evolution of the complete and the reduced system, respectively. Obviously, if the map $\mathcal{U}$ is applied on a valid density matrix, the result must be a valid density matrix as well. This means that the map must be positive and trace-preserving in order to preserve two of the three properties (positive semidefiniteness and trace) of the density matrix. Of course, the Hermitian property must be preserved as well, but as we shall see this is usually trivial. For the map $\mathcal{V}$, however, an additional condition holds, since also every induced map $\mathcal{V}\otimes\hat{I}_{\mathrm{E}}$ must be positive. In other words, even if $\mathcal{V}$ maps a valid reduced density matrix to another, it must be guaranteed that the complete density matrix remains valid as well. If the dimension of the environment E can be chosen arbitrarily, this property is referred to as complete positivity. Fortunately, the Choi-Kraus theorem gives a compact condition whether or not a map is completely positive and trace-preserving (CPTP).

Theorem 2.2.1 (Choi-Kraus) Let $L(\mathcal{H})$ denote the space of square linear operators acting on a Hilbert space $\mathcal{H}$. A map $\mathcal{V}: L(\mathcal{H}) \to L(\mathcal{H})$ is completely positive and trace-preserving (CPTP) if and only if a Choi-Kraus decomposition

$$\mathcal{V}(\hat{\rho}) = \sum_{i=1}^{N^2} \hat{V}_i \hat{\rho} \hat{V}_i^{\dagger} \tag{2.80}$$

exists, where $\hat{\rho} \in L(\mathcal{H})$ and N denotes the dimension of the Hilbert space $\mathcal{H}$, and the Kraus operators $\hat{V}_i \in L(\mathcal{H})$ fulfill

$$\sum_{i=1}^{N^2} \hat{V}_i \hat{V}_i^{\dagger} = \hat{I}. \tag{2.81}$$

Here, $\hat{I} \in L(\mathcal{H})$ denotes the identity operator.

Proof. The proof is omitted for brevity, but can be found e.g., in the text book by Wilde [35]. The original work was published in the papers by Choi [36] and Kraus [37]. ∎

Remark It can be readily seen from Eq. (2.80) that a completely positive and trace-preserving map also preserves the Hermitian property of $\hat{\rho}$.

At this point we know the mathematical properties of the general solution of the master equation without knowledge of the master equation itself. Now it is possible to go along the first path and derive the most general master equation that guarantees a CPTP solution. This approach is purely mathematical and we have to assign a physical meaning to the master equation at a later stage. The only assumptions at this point are that we restrict ourselves to Markovian (or time-local) dynamics, and that the complete system can be prepared initially as $\hat{\rho}(t_0) = \hat{\rho}_{\mathrm{S}}(t_0)\otimes\hat{\rho}_{\mathrm{E}}$. For the sake of brevity, we shall take a shortcut and use the Lindblad-Gorini-Kossakowski-Sudarshan theorem below.

Theorem 2.2.2 (Lindblad-Gorini-Kossakowski-Sudarshan) The solution of a quantum master equation is guaranteed to be a completely positive and trace-preserving map, if the master equation assumes the form

$$\partial_t \hat{\rho}_{\mathrm{S}} = -\mathrm{i}\hbar^{-1}\left[\hat{H}_{\mathrm{S}}, \hat{\rho}_{\mathrm{S}}\right] + \sum_{i,j=1}^{N^2-1} c_{ij}\left(\hat{F}_i \hat{\rho}_{\mathrm{S}} \hat{F}_j^{\dagger} - \frac{1}{2}\hat{F}_j^{\dagger}\hat{F}_i\hat{\rho}_{\mathrm{S}} - \frac{1}{2}\hat{\rho}_{\mathrm{S}}\hat{F}_j^{\dagger}\hat{F}_i\right), \tag{2.82}$$

where the traceless operators $\hat{F}_i$ form an arbitrary orthonormal basis of the space $B(\mathcal{H})$ of bounded operators on the Hilbert space $\mathcal{H}$ with the restriction that $\hat{F}_{N^2}$ is proportional to the identity operator, the coefficients c_{ij} form a positive semi-definite matrix C, and N denotes the dimension of the Hilbert space $\mathcal{H}$.

Proof. The proof is omitted for brevity, but can be found e.g., in the text book by Breuer and Petruccione [34]. The original work was published in the papers by Lindblad [38] and Gorini, Kossakowski and Sudarshan [39]. ∎

By introducing the Lindblad dissipation superoperator

$$\mathcal{D}(\hat{\rho}_\mathrm{S}) = \sum_{i,j=1}^{N^2-1} c_{ij} \left(\hat{F}_i \hat{\rho}_\mathrm{S} \hat{F}_j^\dagger - \frac{1}{2} \hat{F}_j^\dagger \hat{F}_i \hat{\rho}_\mathrm{S} - \frac{1}{2} \hat{\rho}_\mathrm{S} \hat{F}_j^\dagger \hat{F}_i \right), \tag{2.83}$$

we can rewrite Eq. (2.82) as the extension of the Liouville-von Neumann equation (2.65). As already anticipated in Section 2.2.4, an extra dissipation term has appeared. This is the consequence of the transition from a closed system, in which the probability values p_k cannot be affected and the dynamical behaviour is governed by unitary evolution, to an open system, in which the probabilities can vary due to interactions with the environment. The resulting form

$$\partial_t \hat{\rho}_\mathrm{S} = \mathcal{L}(\hat{\rho}_\mathrm{S}) + \mathcal{D}(\hat{\rho}_\mathrm{S}) \tag{2.84}$$

of the master equation shall be referred to as the Lindblad equation.

Alternatively, different forms of the quantum master equation can be derived microscopically by starting from Eq. (2.77) and exploiting different assumptions and approximations. For example, the Markovian approximation, the Born approximation (weak coupling limit), and the secular approximation are required in order to derive a master equation of Lindblad form microscopically. This path has a stronger focus on the physical processes between the system under consideration and the environment. It should be noted that some forms, such as the Redfield equation, do not guarantee a CPTP solution and therefore might violate the mathematical properties of the reduced density matrix. Nevertheless, they can still be useful in practice when used with care.

2.3 Basic theory of semiconductors and semiconductor heterostructures

The behavior of semiconductors can be explained basing on the postulates of quantum mechanics. However, this is a formidable application of quantum mechanics and requires significant modeling effort and auxiliary concepts. Once more it should be noted that a complete treatment of semiconductor physics is beyond the scope of this thesis. Instead, we derive in this section the subset of concepts that is required to understand the state of the art in modeling semiconductor lasers, in particular QCLs. To this end, we shall employ several existing models and approximations that are commonly used in semiconductor physic, but omit their detailed introduction for the sake of brevity. The interested reader is referred to the many excellent books on the subject, e.g., the introduction to solid state physics by Ashcroft and Mermin [40], the general introduction to semiconductor physics by Yu and Cardona [41], or the books on low-dimensional semiconductor structures and semiconductor heterostructures by Davies [42] and Bastard [43], respectively.

2.3.1 Band structure of semiconductors

Let us first consider a free electron, i.e., an electron that is not confined by any potential and is free to move. In this case, the Hamiltonian $\hat{H}_0 = \hat{\boldsymbol{p}}^2/(2m_0)$, where m_0 denotes the electron mass, is time-independent and the stationary Schrödinger equation (2.46), which then reads

$$-\frac{\hbar^2 \nabla^2}{2m_0} \Psi(\boldsymbol{r}) = \epsilon \Psi(\boldsymbol{r}), \tag{2.85}$$

can be used to determine the eigenenergies of the free electron. By plugging in a suitable ansatz $\Psi(\boldsymbol{r}) = A \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})$, where A is an arbitrary amplitude of a wave with the wave vector $\boldsymbol{k}$, it becomes apparent that the eigenenergies are continuous and form a dispersion relation

$$\epsilon(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m_0}, \tag{2.86}$$

which means that the electrons may assume an arbitrary energy value.

The situation is significantly more complex when electrons in a semiconductor material are considered. Solid state semiconductors consist of a crystal lattice in which the vertices are occupied by the atom cores. The electrons are usually divided in valence electrons and core electrons, where the latter are usually assumed to be localized around the atom cores. The energy of the system is the sum of the kinetic energy of the atom cores and electrons, respectively, and the potential energy that stems from the Coulomb forces (cf. Eq. (2.1), but neglecting the magnetic field $\boldsymbol{B} = 0$) between the different atom cores, between the atom cores and the electrons, as well as between the different electrons. This many-particle approach is exceedingly complex and usually simplified with the following approximations. First, it is assumed that only the valence electrons contribute to the dynamic behavior of the semiconductor and the core electrons are treated in combination with the atom core. Then, the Born-Oppenheimer (or adiabatic) approximation exploits the difference in inertia between the atom cores (including the core electrons) and the electrons. As a consequence, the atom cores experience a time-averaged adiabatic potential that stem from the electrons, and from the point of view of the electrons the atom cores are essentially stationary. Finally, the interactions between the individual electrons are averaged by virtue of the mean-field approximation, which assumes that every electron experiences the same potential $V(\boldsymbol{r})$. In the resulting description, a single exemplary electron can be described using the time-independent Schrödinger equation, in which of course the potential is assumed to be time-independent. Any variations of the potential in time can be included in a later step using perturbation theory to describe the transitions of the electrons between their quantum mechanical states [41].

Even without knowledge of the exact form of the potential it is possible to extract a significant characteristic of semiconductors from the single-particle description. To this end, we assume that the semiconductor is an infinite perfect crystal lattice with the period $\boldsymbol{R}$, i.e., $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R})$, and use Bloch's theorem.

Theorem 2.3.1 (Bloch) The eigenstates of a single-particle Hamiltonian $\hat{H}_0 = \hat{\boldsymbol{p}}^2/(2m) + V(\boldsymbol{r})$, where the potential $V(\boldsymbol{r}) = V(\boldsymbol{r} + \boldsymbol{R})$ is periodic with the period $\boldsymbol{R}$, can be chosen to have the form of a plane wave multiplied with a function with the same periodicity. In other words, the eigenstates assume the form

$$\Psi_{n,\boldsymbol{k}}(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, u_{n,\boldsymbol{k}}(\boldsymbol{r}), \tag{2.87}$$

where $u_{n,\boldsymbol{k}}(\boldsymbol{r}) = u_{n,\boldsymbol{k}}(\boldsymbol{r} + \boldsymbol{R})$ holds.

Proof. The proof is omitted for brevity, but can be found e.g., in the text book by Ashcroft and Mermin [40]. The original work was published in the papers by Floquet [44] (for the one-dimensional case) and Bloch [45] (for the generalized three-dimensional case). ■

Remark For a given wave vector $\boldsymbol{k}$ there may be more than one solution to the stationary Schrödinger equation, i.e., more than one eigenstate. The band index n has been introduced to distinguish between the different families of solutions. Each family of solutions features a continuous function $\epsilon_n(\boldsymbol{k})$ for the corresponding eigenenergy values [40].

The notion of bands, i.e., states with a continuous relation between energy and wave vector (dispersion relation), is a significant consequence of Bloch's theorem. Each band provides a continuous range of allowed energy values, but introduces the possibility that between two bands there could be a gap. Indeed, such a band gap exists in the band structure of insulator and semiconductor materials, and constitutes a range of energy values that the electrons cannot assume. Furthermore, it separates the valence band (the highest range of energy values that are occupied by electrons at absolute zero temperature) from the conduction band (the lowest range of energy values that are vacant at absolute zero temperature). In a wider definition, all bands below the band gap are referred to as valence bands, and all bands above the band gap are called conduction bands. The energy of the electrons in the conduction band exceeds the Fermi energy, which allows them to move and, as a consequence, contribute to electric conduction. Therefore, whether or not a band gap arises, and what energy range it covers, defines whether the material under discussion is a conductor (no band gap), an insulator (large band gap), or semiconductor (small band gap). This information can only be obtained from the band structure, which requires exact knowledge of the dispersion relation of all bands. The dispersion relation, in turn, requires the knowledge of the exact form of the potential $V(\boldsymbol{r})$.

2.3.2 The effective mass approximation

Various methods exist that are able to determine the band structure of a certain semiconductor material [40], [41]. For example, plugging in the eigenstates Eq. (2.87) into the stationary Schrödinger equation (2.46) yields the eigenvalue equation

$$\left[\frac{\hat{\boldsymbol{p}}^2}{2m_0} + V(\boldsymbol{r}) + \frac{\hbar}{m_0}\boldsymbol{k}\cdot\hat{\boldsymbol{p}} + \frac{\hbar^2\boldsymbol{k}^2}{2m_0}\right] u_{n,\boldsymbol{k}}(\boldsymbol{r}) = \epsilon_n(\boldsymbol{k})\, u_{n,\boldsymbol{k}}(\boldsymbol{r}) \tag{2.88}$$

for the periodic envelope function $u_{n,\boldsymbol{k}}(\boldsymbol{r})$. This equation serves as starting point for the k.p method (named after the third term of the Hamiltonian operator) and its variations. The k.p method solves the eigenvalue equation for $\boldsymbol{k} = 0$, which yields the states $u_{n,0}(\boldsymbol{r})$. The terms in the Hamiltonian that contain the wave vector are subsequently considered using the states as basis and time-independent perturbation theory. Usually, the center $\boldsymbol{k} = 0$ is assumed to be an extremum of the dispersion relation $\epsilon_n(\boldsymbol{k})$ and only small deviations of the wave vector from the center are considered. In this case, second order perturbation theory is sufficient to approximate the dispersion relation with the expression

$$\epsilon_n(\boldsymbol{k}) \approx \epsilon_n(0) + \frac{\hbar^2\boldsymbol{k}^2}{2m_0}\left\{1 + \frac{2}{m_0}\sum_{m,m\neq n}\frac{\left|\langle u_{m,0}|\boldsymbol{k}\cdot\hat{\boldsymbol{p}}|u_{n,0}\rangle\right|^2}{\boldsymbol{k}^2\left[\epsilon_n(\boldsymbol{0}) - \epsilon_m(\boldsymbol{0})\right]}\right\}. \tag{2.89}$$

In principle, however, the method is able to determine the band structure globally (i.e., without the restriction of the wave vector to the region around the center) provided that higher-order perturbation terms are included [42]. We compare Eq. (2.89) to the dispersion relation for the free electron in Eq. (2.86) and find that they only differ by the offset $\epsilon_n(0)$ in energy and the complex expression in the curly braces. Indeed, the dispersion relation of a certain band around the center $\boldsymbol{k} = 0$ can frequently be approximated with a parabola similar to Eq. (2.86), where the mass is replaced with the effective mass of the band under consideration.

Assuming that we already know the band structure of a certain semiconductor material (by means of the k.p method or one of the various alternatives), we can derive the effective mass of a certain band by expanding the corresponding dispersion relation

$$\epsilon_n(\boldsymbol{k}) = \epsilon_n(0) + \sum_i \partial_{k_i}\epsilon_n(\boldsymbol{k})\big|_{\boldsymbol{k}=0}\, k_i + \frac{1}{2}\sum_{i,j}\partial_{k_i}\partial_{k_j}\epsilon_n(\boldsymbol{k})\big|_{\boldsymbol{k}=0}\, k_i k_j + O(\boldsymbol{k}^3) \tag{2.90}$$

as Taylor series at the center, where the k_i denote the components of the wave vector and the indices $i, j \in \{x, y, z\}$ represent the directions in 3D. Again, we assume that the center $\boldsymbol{k} = 0$ is an extremum and, as a consequence, the first-order derivatives $\partial_{k_i}\epsilon_n(\boldsymbol{k})\big|_{\boldsymbol{k}=0} = 0$ in all directions. We truncate the series after the second order term and compare the result

$$\epsilon_n(\boldsymbol{k}) \approx \epsilon_n(0) + \frac{1}{2}\boldsymbol{k}^{\mathrm{T}}\mathbf{H}_{\epsilon_n}(0)\boldsymbol{k}, \tag{2.91}$$

where $\mathbf{H}_{\epsilon_n}$ is the Hessian of the dispersion relation $\epsilon_n(\boldsymbol{k})$, to Eq. (2.86). Ignoring the offset $\epsilon_n(0)$ for a moment, we can readily see that they correspond to each other provided that the electron mass m_0 is replaced with the effective mass tensor

$$\mathbf{M}_n^* = \hbar^2\left[\mathbf{H}_{\epsilon_n}(0)\right]^{-1}. \tag{2.92}$$

With the help of this artificial mass (which is a tensor in general, and even may be negative in certain cases) we can treat electrons in a semiconductor material as we would treat free quantum mechanical particles. In the next section, we shall discuss the application of this concept to the modeling of semiconductor heterostructures, in which we also use the (previously ignored) offset $\epsilon_n(0)$.

2.3.3 Modeling of semiconductor heterostructures

Over the last decades, many semiconductor devices have been developed that exploit the physical properties of heterojunctions, such as the high electron mobility transistor (HEMT) or the quantum well laser. A

heterojunction is defined as the interface between two different semiconductor materials (in contrast to homojunctions, such as the well-known p-n junction, which are the consequence of differences in doping). The combination of multiple heterojunctions are referred to as semiconductor heterostructure. The QCL gain medium is a prime example of such a heterostructure as it consists of a stack of thin layers with different semiconductor materials. Therefore, we derive in the following the basic concepts required to model the QCL gain medium.

To this end, we again make some basic assumptions that simplify the resulting model significantly. From the band structures of the different semiconductor materials we only consider the bottom conduction band as relevant. As a consequence, interband transitions (e.g., from the conduction band to one of the valence bands) are ignored. Instead, the transitions between states within the conduction band, which are usually referred to as subbands, are considered. Additionally, we continue to assume that the conduction band of each semiconductor material in use can be represented by a parabolic dispersion relation. This is accompanied by the assumption that only a small range of wave vectors around the center $\boldsymbol{k} = 0$ contribute to the allowed states in a semiconductor heterostructure.

With these assumptions in mind, we aim to derive a simplified Schrödinger equation that yields the allowed states in a heterostructure. To this end, we consider the state

$$\Psi(\boldsymbol{r}) = \int c(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, u_{\boldsymbol{k}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \approx u_0(\boldsymbol{r}) \int c(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \tag{2.93}$$

as linear superposition of Bloch states with the coefficients c, where the integral reflects the continuous nature of the wave vector $\boldsymbol{k}$ and the band index n has been dropped as only the conduction band is considered. Since we assume that only a small range of wave vectors contribute to Ψ, the function $u_{\boldsymbol{k}} \approx u_0$ is approximately constant and can be extracted from the integral. This step is commonly referred to as envelope function approximation (EFA). The remaining integral can be interpreted as inverse Fourier transform of the coefficients $c(\boldsymbol{k})$ from the pseudo-spectral $\boldsymbol{k}$-space, which yields a certain envelope function

$$\psi(\boldsymbol{r}) = \int c(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, \mathrm{d}\boldsymbol{k}. \tag{2.94}$$

As the state $\Psi(\boldsymbol{r})$ must fulfill the stationary Schrödinger equation $\hat{H}_0\Psi(\boldsymbol{r}) = \epsilon\Psi(\boldsymbol{r})$, we investigate the effect

$$\hat{H}_0\Psi(\boldsymbol{r}) = \int c(\boldsymbol{k})\, \hat{H}_0 \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, u_{\boldsymbol{k}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} = \int c(\boldsymbol{k})\, \epsilon(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, u_{\boldsymbol{k}}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \tag{2.95}$$

of the Hamiltonian $\hat{H}_0 = \hat{\boldsymbol{p}}^2/(2m) + V(\boldsymbol{r})$ on the state $\Psi(\boldsymbol{r})$. By invoking the envelope function approximation, this term can be simplified to

$$\hat{H}_0\Psi(\boldsymbol{r}) \approx u_0(\boldsymbol{r}) \int c(\boldsymbol{k})\, \epsilon(\boldsymbol{k}) \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, \mathrm{d}\boldsymbol{k}. \tag{2.96}$$

At this point, Eq. (2.91) can be plugged in for the dispersion relation $\epsilon(\boldsymbol{k})$, which yields

$$\hat{H}_0\Psi(\boldsymbol{r}) \approx u_0(\boldsymbol{r}) \left[\epsilon(0)\psi(\boldsymbol{r}) + \frac{1}{2} \int c(\boldsymbol{k})\, \boldsymbol{k}^{\mathrm{T}} \mathbf{H}_\epsilon(0) \boldsymbol{k} \exp(\mathrm{i}\boldsymbol{k}\boldsymbol{r})\, \mathrm{d}\boldsymbol{k} \right]. \tag{2.97}$$

We note that the Hessian $\mathbf{H}_\epsilon$ does not depend on the wave vector and can be treated as constant. Then, from basic Fourier theory we know that a multiplication with $\boldsymbol{k}$ in $\boldsymbol{k}$-space corresponds to a differentiation operation $-\mathrm{i}\nabla$ in regular space and we can treat the integral as inverse Fourier transformation. After the transformation, the effect of the Hamiltonian reads

$$\hat{H}_0\Psi(\boldsymbol{r}) \approx u_0(\boldsymbol{r}) \left[\epsilon(0) - \frac{1}{2} \nabla \mathbf{H}_\epsilon(0) \nabla \right] \psi(\boldsymbol{r}). \tag{2.98}$$

Finally, we find that in the envelope function approximation, the stationary Schrödinger equation

$$u_0(\boldsymbol{r}) \left[\epsilon(0) - \frac{1}{2} \nabla \mathbf{H}_\epsilon(0) \nabla \right] \psi(\boldsymbol{r}) = \epsilon u_0(\boldsymbol{r}) \psi(\boldsymbol{r}) \tag{2.99}$$

for the state $\Psi(\boldsymbol{r})$ can be reduced to the equation

$$\left[-\frac{\hbar^2}{2}\nabla\left(\mathbf{M}_n^*\right)^{-1}\nabla - V_\mathrm{c}\right]\psi\left(\boldsymbol{r}\right) = \epsilon\psi\left(\boldsymbol{r}\right) \tag{2.100}$$

for the envelope $\psi\left(\boldsymbol{r}\right)$, where we plugged in the effective mass tensor and introduced the potential V_c to replace $\epsilon(0)$.

Equation (2.100) can now be applied to a semiconductor heterostructure by allowing the effective mass tensor and the potential V_c to vary in growth direction z. The resulting function $V_\mathrm{c}(z)$ is commonly referred to as conduction band profile and features multiple quantum wells and barriers. An example of such a profile has already been depicted in Fig. 1.3. For the effective mass, we make two additional assumptions. Since the effective mass tensor is symmetric, it can be diagonalized by means of a suitable change of basis. In other words, there exists a principal axis system in which the tensor is diagonal but it is not guaranteed that this axis system coincides with the coordinate system of the device to be modeled. Nevertheless, a diagonal effective mass tensor is frequently assumed in the modeling of heterostructures. Additionally, in the Ben Daniel-Duke model it is assumed that the main diagonal entries $m_{xx}^* = m_{yy}^* = m^{||}$ related to the in-plane directions x and y are equal, where the in-plane effective mass $m^{||}$ is introduced for convenience. The remaining diagonal entry $m_{zz}^* = m^*$ is referred to as the effective mass in growth direction z. With these assumptions in mind, we can write Eq. (2.100) as

$$\left\{-\frac{\hbar^2}{2}\left[\frac{1}{m^{||}(z)}\left(\partial_x^2 + \partial_y^2\right) + \partial_z\frac{1}{m^*(z)}\partial_z\right] + V_\mathrm{c}(z)\right\}\psi(\boldsymbol{r}) = \epsilon\psi(\boldsymbol{r}). \tag{2.101}$$

Since the conduction band profile and the effective masses only vary in growth direction, we can make the ansatz

$$\psi(\boldsymbol{r}) = S^{-1/2}\psi_{\boldsymbol{k}_\mathrm{T}}(z)\exp\left(\mathrm{i}k_{\mathrm{T},x}x + \mathrm{i}k_{\mathrm{T},y}y\right), \tag{2.102}$$

where $\boldsymbol{k}_\mathrm{T} = [k_{\mathrm{T},x}, k_{\mathrm{T},y}]^\mathrm{T}$ is the in-plane wave vector, and the in-plane cross section area S ensures the normalization $\int|\psi_{\boldsymbol{k}_\mathrm{T}}(z)|^2\mathrm{d}z = 1$. Plugging in this ansatz into Eq. (2.101) yields the Ben Daniel-Duke model

$$\left\{\frac{\hbar^2}{2}\left[\frac{k_{\mathrm{T},x}^2 + k_{\mathrm{T},y}^2}{m^{||}(z)} + \partial_z\frac{1}{m^*(z)}\partial_z\right] + V_\mathrm{c}(z)\right\}\psi_{\boldsymbol{k}_\mathrm{T}}(z) = \epsilon_{\boldsymbol{k}_\mathrm{T}}\psi_{\boldsymbol{k}_\mathrm{T}}(z), \tag{2.103}$$

which can be simplified to a one-dimensional stationary Schrödinger equation

$$\left[-\frac{\hbar^2}{2}\partial_z\frac{1}{m^*(z)}\partial_z + V_\mathrm{c}(z)\right]\psi(z) = \epsilon\psi(z) \tag{2.104}$$

by neglecting the spatial dependence of the in-plane effective mass. For the kinetic energy, the relation $\epsilon_{\boldsymbol{k}_\mathrm{T}} = \epsilon + \hbar^2\boldsymbol{k}_\mathrm{T}^2/(2m^{||})$ holds in this case.

Equation (2.104) provides a simple model that is widely used in the simulation of semiconductor heterostructure devices, such as the QCL. It should be noted that it must be complemented by appropriate boundary conditions. Apart from the obvious condition $\psi(z_0^+) = \psi(z_0^-)$ for a continuous wave function (where $z_0^\pm$ represents the two sides of a heterojunction at the position z_0), a second boundary condition can be derived by integrating over Eq. (2.104). Since

$$\begin{aligned} 0 &= \int_{z_0-\alpha}^{z_0+\alpha}\left[-\frac{\hbar^2}{2}\partial_z\frac{1}{m^*(z)}\partial_z + V_\mathrm{c}(z) - \epsilon\right]\psi(z)\mathrm{d}z \\ &= -\frac{\hbar^2}{2}\frac{1}{m^*(z)}\partial_z\psi(z)\Big|_{z_0-\alpha}^{z_0+\alpha} + \int_{z_0-\alpha}^{z_0+\alpha}\left[V_\mathrm{c}(z) - \epsilon\right]\psi(z)\mathrm{d}z \end{aligned} \tag{2.105}$$

must hold, we get for $\alpha \to 0$

$$0 = -\frac{\hbar^2}{2}\frac{1}{m^*(z)}\partial_z\psi(z)\Big|_{z_0-\alpha}^{z_0+\alpha} \tag{2.106}$$

and, as a consequence, the boundary condition $m^*(z_0^+)\partial_z\psi(z_0^-) = m^*(z_0^-)\partial_z\psi(z_0^+)$.

2.4 Summary

The main product of any laser is electromagnetic radiation that can be described with the laws of classical electrodynamics, namely the equation for the Lorentz force and Maxwell's equations. Starting from the microscopic formulation of Maxwell's equations, different alternative descriptions of these equations have been derived. For example, the macroscopic formulation is usually more suitable in the scope of device simulation. The source terms in both formulations, such as the current density and the charge density, as well as their variations have been discussed in detail. Additionally, the time-harmonic form of Maxwell's equations and the corresponding complex relative permittivity have been introduced. The potential formulation exploits the redundancy of Maxwell's equations and uses the electromagnetic scalar and vector potentials to describe the corresponding physical processes without any loss of generality. The gauge invariance allows further reduction of complexity in the equations, yielding a very compact notation, which is commonly used in physics.

The gain medium of the QCL, on the other hand, cannot be described using classical physics. Therefore, we have reviewed the basic postulates of quantum mechanics and their implications. The state of a system is described by a state vector, and operators that act on the state vector correspond to physical quantities. We have discussed the properties of the operators and their relation to the macroscopic, classical quantities. The time evolution of the quantum mechanical states can be described with the Schrödinger equation, in which the Hamiltonian operator plays a pivotal role. For a the time-independent Hamiltonian, the concept of eigenenergies becomes apparent. The eigenenergies can be related to the discrete energy levels already mentioned in Chapter 1. In general, however, the Hamiltonian is time-dependent, for example for a quantum mechanical system in a classical electromagnetic field. This example has been introduced as preparation for the interaction between the gain medium and the electromagnetic radiation. Since a description using state vectors is only convenient for a system with few particles, we have identified the density operator as central concept to treat many-particle systems. When using the density operator, the time evolution of the system is governed by the Liouville-von Neumann equation. Often it makes sense to divide the system in a reduced system of interest and an environment in order to reduce the complexity. Therefore, we have introduced the notion of composite systems and quantum master equations, which describe the time evolution of the reduced density operator.

The behavior of semiconductors can be described using the basic postulates of quantum mechanics, but this is a formidable task that is beyond the scope of this chapter. However, we have derived several basic concepts that are common in semiconductor theory, such as the band structure of semiconductor materials. Using the concept of the effective mass tensor, we have related the movement of an electron in the crystal lattice of a semiconductor to that of a free electron. Finally, using the envelope function approximation a very compact model for the states in a semiconductor heterostructure has been derived. The path from the basic postulates of quantum mechanics to this compact model leads over several assumptions and approximations, which may have been discussed only briefly but provide a complete overview. This overview will serve as basis for the discussion of quantum cascade laser models in the next chapter.

3 Modeling the dynamics of quantum cascade lasers

> This is awesome! If only real life was in 3D.
> — *Bart Simpson*

> I have the solution, but it works only in the case of spherical cows in a vacuum!
> — *The oldest modeling joke*

Building on the general theoretical basis in Chapter 2, we advance to the modeling of quantum cascade lasers. In alignment with the division between gain medium and resonator cavity, we start with a review of different modeling approaches for the electron dynamics in the gain medium in Section 3.1. Here, we exploit a feature that all of them have in common – namely, the separation of the quantum mechanical states into envelope wave functions that depend only on the growth direction z and components that depend on the in-plane coordinates x and y. This separation facilitates the treatment of the gain medium as collection of quantum mechanical systems that occupy a certain position in the x-y-plane and span the complete active region in z direction. The dynamical behavior of this collection of quantum mechanical systems can then be described with a master equation of Lindblad form (cf. Eq. (2.84)), where the Hamiltonian of the reduced system describes the physical effects in growth direction and the dissipation term represents the in-plane perturbation effects. Then, we address the coupling between the electromagnetic field in the laser and the time evolution of the quantum mechanical systems. Here, we can safely assume that the dimensions of the quantum mechanical systems are significantly smaller than the occurring wavelengths of the optical field. This assumptions is commonly referred to as dipole approximation. We continue in Section 3.2 with the dynamics of an electromagnetic field in the resonator cavity, where we establish the assumption that the total electric field

$$\boldsymbol{E} = \underbrace{-\nabla\varphi}_{\boldsymbol{E}_\mathrm{B}} \underbrace{-\partial_t \boldsymbol{A}}_{\boldsymbol{E}_\mathrm{O}} \tag{3.1}$$

consists of a static bias field $\boldsymbol{E}_\mathrm{B}$ and a dynamic optical field $\boldsymbol{E}_\mathrm{O}$, which are directly related to the scalar potential φ and vector potential $\boldsymbol{A}$, respectively. Then, we exploit the fact that the resonator geometry is constant in propagation direction and derive a one-dimensional propagation model, which reduces the overall complexity significantly. As the different modeling approaches feature (partial) differential equations that cannot be solved analytically in general, Section 3.3 gives an overview of simulation software packages that solve them numerically. As a conclusion, we identify the Maxwell-Bloch equations as key model for the quantum cascade laser dynamics that will be in the focus for the remainder of the thesis at hand.

3.1 Electron dynamics in the gain medium

Over the last decades, many different approaches to model the charge carrier transport in quantum cascade lasers have been presented. They vary in complexity and accuracy, and can be divided into four categories: rate equations, ensemble Monte Carlo approaches, density matrix methods, and techniques that base on non-equilibrium Green's functions (NEGF). All approaches share the common assumption that the states of the electrons

$$\Psi_i(\boldsymbol{r}) = S^{-1/2}\psi_{i,\boldsymbol{k}}(z)\exp\left(\mathrm{i}k_x x + \mathrm{i}k_y y\right) \tag{3.2}$$

can be separated into the wave function envelope $\psi_{i,\boldsymbol{k}}(z)$ that varies only in growth direction and a plane wave with the in-plane wave vector $\boldsymbol{k} = [k_x, k_y]^\mathrm{T}$ (cf. Eq. (2.102) in Section 2.3, where we introduced the in-plane cross section area S to ensure normalization) [9]. This assumption has two important consequences that simplify the resulting models significantly. First, the in-plane wave ansatz allows the extraction of a single quantum mechanical system that spans the complete active region in z direction from the x-y-plane. The dynamical behavior of a single system can then be discussed, and at a later stage, the complete active region can be considered as collection of single quantum systems. Second, during the discussion of the dynamical behavior we can separate the physical effects in growth direction, such as the formation of discrete basis states due to the semiconductor heterostructure (represented by the wave function envelopes $\psi_{i,\boldsymbol{k}}(z)$) and radiative transitions between the states, from the in-plane effects that can be roughly said to define the non-radiative transitions (scattering) between the states. The resulting picture has proven to be compact and intuitive, yet sufficiently accurate.

Using this picture, the least complex group of models considers the occupation of the basis states in simple rate equations, similar to those used for conventional lasers [10]. The values of the scattering rates are chosen phenomenologically. One-dimensional density matrix methods extend the simple rate equations by also considering the coherence terms between the basis states. In other words, the time evolution of the full density matrix is determined rather than that of the main diagonal alone. The resulting equations still depend on empirically determined parameters. ensemble Monte Carlo (EMC) methods [9], [46], [47], on the other hand, take the in-plane movement of charge carriers into account and provide a self-consistent method to determine the scattering rates. However, they neglect the coherence terms and other quantum effects and are therefore considered semi-classical. Three-dimensional density matrix methods [48]–[50] and techniques that base on NEGF [51]–[53] are both fully quantum mechanical descriptions that consider the movement of the charge carriers in all directions. Hence, they constitute the most accurate group of methods, which comes at the cost of huge computational complexity. Finally, as a trade-off between accuracy and complexity, hybrid models have been presented. A prime example of this group is the density matrix ensemble Monte Carlo (dmEMC) technique [54] that introduces quantum corrections to semi-classical Monte Carlo methods, thereby combining the efficiency of the Monte Carlo framework with the accuracy of fully quantum mechanical models.

Three-dimensional density matrix techniques as well as EMC and NEGF methods have been used successfully for precise stationary simulations, such as the accurate description of the optical gain of the active region. However, their computational load inhibits their usage in dynamical simulations, where the quantum mechanical description has to be updated frequently due to the varying electromagnetic field. Rate equations, on the other hand, have been used for dynamical simulations (for example, to study the gain recovery dynamics [55] of quantum cascade lasers), but they do not include the coherence terms. Instead, one-dimensional density matrix methods have been employed frequently for dynamical simulations as they provide quantum mechanical treatment of the electron dynamics with reduced complexity. The density matrix in these methods can be considered as the reduced density matrix of the system, where the environment contains the in-plane degrees of freedom. Consequently, it is possible to derive a master equation for the reduced density matrix, which corresponds to the optical Bloch equations when written in elementwise form [32]. Together with Maxwell's equations for the propagating electromagnetic field, the equations are referred to as the Maxwell-Bloch equations.

In the following, we derive the master equation for the reduced density matrix. To this end, we select a suitable set of basis states, discuss the possibilities to determine the dissipation superoperator in a self-consistent way, and finally model the interaction between the quantum mechanical systems and the electromagnetic field.

3.1.1 Basis states of the electrons

There are many different ways to obtain the electronic states in quantum cascade lasers. For example, in many NEGF approaches the Wannier(-Stark) states are simply assumed as basis [51]. In order to provide more accurate results, however, the basis states are often determined numerically using models of different complexity. Frequently, the k.p method is used as basis to derive those models [56]. The k.p method was

already mentioned in Section 2.3 as one possible way to determine the band structure of a semiconductor. From this starting point, different models can be derived which are distinguished by the considered energy bands. For example, the Kane model and the Luttinger-Kohn model consider three valence bands and the conduction band, which results in a description of four or eight bands in total, depending on whether the electron spin degeneracy is taken into account or not [42]. With those models it is possible to skip the assumption of a single parabolic band, where the wave vector should be close to the extremum of the band. Thereby, non-parabolic effects such as strain between the heterostructure layers can be taken into account. The Ben Daniel-Duke model, on the other hand, considers only the conduction band. Particularly for terahertz QCLs this is a reasonable approximation seeing that the energy differences between the resulting basis states are significantly smaller than the gap between the conduction band and the valence bands [41]. Since those energy differences are larger in mid-infrared QCLs, more elaborated models may be used, such as the three-band approach in [50]. Nevertheless, the Ben Daniel-Duke model has proven to be a valuable tool even for mid-infrared QCLs [9]. Therefore, we discuss this model in more detail in the remainder of this section to provide one possible way to determine the basis states of the electrons in the QCL gain medium.

We recall that in the scope of the Ben Daniel-Duke model (and by assuming a constant in-plane effective mass $m^{||}$) we can describe the states in a semiconductor heterostructure with the one-dimensional stationary Schrödinger equation (2.104), which yields the wave function envelopes. The complete wave function can then be put together according to Eq. (3.2). Therefore, we define a similar Schrödinger equation

$$\left[-\frac{\hbar^2}{2}\partial_z \frac{1}{m^*(z)}\partial_z + V(z)\right]\psi_i(z) = \epsilon_i \psi_i(z) \tag{3.3}$$

for the wave function envelopes $\psi_i(z)$, where the potential energy V contains the conduction band profile V_c. At this point, we note that we need to take the bias field into account, which we defined as $\boldsymbol{E}_\mathrm{B} = -\nabla\varphi$ in Eq. (3.1). According to Eq. (2.48), the effect of the scalar potential φ on the Hamiltonian can be expressed by an extra term in the potential $V(z) = V_\mathrm{c}(z) - e\varphi(z)$, where e is the elementary charge and we assume that the effect of the bias voltage is independent of the position in the x-y-plane. The effect of the optical field, on the other hand, is usually neglected during the calculation of the basis states and considered as a time-dependent perturbation at a later point. Therefore, we can write for the time-independent Hamiltonian

$$\hat{H}_0 = -\frac{\hbar^2}{2}\partial_z \frac{1}{m^*(z)}\partial_z + V_\mathrm{c}(z) - e\varphi(z). \tag{3.4}$$

In principle, Eq. (3.3) can be solved numerically in the current form with appropriate boundary conditions. However, there are some variations that should be mentioned briefly. Although non-parabolicity effects are neglected by default in the Ben Daniel-Duke model, it is possible to take them into account by means of an energy dependent effective mass $m^*(z, \epsilon)$, whose form can be determined beforehand with the k.p method. The drawback of this approach is that the resulting Hamiltonian $\hat{H}_0$ is not Hermitian in this case and the resulting basis states are in general not orthogonal [9]. Another variation of the model above considers only one period of the heterostructure and determines the basis states in a tight-binding approach. The resulting tight-binding (localized) basis has proven to be advantageous, e.g., to account for tunneling between thick barriers in the heterostructure in semiclassical models [54]. Finally, there are different models for the effect of the bias voltage or the scalar potential φ, respectively. In the easiest case, a linear scalar potential $\varphi(z) = U_\mathrm{B} z/L_z$ is assumed, where L_z is the size of the heterostructure in growth direction, and U_B is the applied bias voltage. A more advanced model shall be presented in Section 3.2, where the scalar potential is determined by solving the Poisson equation. The resulting Schrödinger-Poisson model is able to take e.g., space charge effects due to doping into account and is widely used in QCL simulations.

3.1.2 Master equation and dissipation superoperator

Independent of how they have been determined, we can use the basis states discussed in the section before (or a subset thereof) to define the density matrix of a reduced quantum mechanical system. This system contains all states that are relevant for the quantum cascade laser operation, where each state – as pointed out

before – considers only the growth direction. Assuming that the in-plane dynamics have only little impact on the basis states and vice-versa (weak coupling limit), and that the physical processes have no memory (Markovian approximation), the dynamics of the reduced system can be described by a master equation

$$\partial_t \hat{\rho} = -\mathrm{i}\hbar^{-1}\left[\hat{H}, \hat{\rho}\right] + \mathcal{D}(\hat{\rho}) \tag{3.5}$$

of Lindblad type (cf. Section 2.2.6). For the remainder of this thesis, we drop the index from the reduced density matrix $\hat{\rho}_S =: \hat{\rho}$ for brevity. The Hamiltonian $\hat{H}$ contains the Hamiltonian $\hat{H}_0$ in Eq. (3.4) and shall be extended later by a term that accounts for the interaction with the optical field, as already envisaged in the previous section. Often, the basis states are eigenstates of the Hamiltonian and $\hat{H}_0$ can be represented by a diagonal matrix, where the main diagonal contains the eigenenergies. However, we note that this is not necessarily the case. For example, when the tight-binding approach is used, off-diagonal terms may appear.

The dissipation superoperator $\mathcal{D}$ in Eq. (2.83) gives a very general description of a master equation whose solution is guaranteed to be completely positive and trace-preserving. Now we need to assign the general description a physical meaning. We recall that the operators $\hat{F}_k$ in Eq. (2.83) are traceless and constitute an orthonormal basis of the space of bounded operators on the Hilbert space under consideration. With the basis states determined in the section before, the bounded operators are $N \times N$ Hermitian matrices, where N is the number of chosen basis states. Those matrices have $N(N-1)$ degrees of freedom in the off-diagonal elements, as the Hermitian property allows for complex numbers, but requires that the lower triangle elements are the complex conjugate of the upper triangle elements. The N real main diagonal elements add further $N-1$ degrees of freedom, since one degree is already taken into account by the basis operator $\hat{F}_{N^2}$. A suitable choice of basis operators is presented in [57] and consists of $N(N-1)$ off-diagonal operators of the form $\hat{F}_k = |i\rangle\langle j|$, where the indices $i, j \in [1; N], i \neq j$ are mapped to the index $k \in [N; N^2 - 1]$, and $N-1$ diagonal operators of the form

$$\hat{F}_k = \frac{1}{\sqrt{k(k+1)}}\left(\sum_{s=1}^{k} |s\rangle\langle s| - k\,|k+1\rangle\langle k+1|\right), \tag{3.6}$$

where $k \in [1; N-1]$. The latter group of operators can be related to the Pauli matrices and Gell-Mann matrices (for $N = 2$ and $N = 3$, respectively). Using these basis operators, we need to determine the coefficients c_{ij} in Eq. (2.83) and verify that the resulting matrix is positive semi-definite.

We begin with the off-diagonal operators and assume that the coefficients $c_{ij} = 0$ for $i, j \geq N, i \neq j$. Although this assumption may seem to be overly strict, the resulting contribution

$$\mathcal{D}_{\mathrm{relax}}(\hat{\rho}) = \sum_{k=N}^{N^2-1} c_{kk}\left(\hat{F}_k\hat{\rho}\hat{F}_k^\dagger - \frac{1}{2}\hat{F}_k^\dagger\hat{F}_k\hat{\rho} - \frac{1}{2}\hat{\rho}\hat{F}_k^\dagger\hat{F}_k\right), \tag{3.7}$$

to the dissipation superoperator provides a physically meaningful description of the population relaxation processes. By plugging in the off-diagonal basis operators, we find that each $\hat{F}_k$ generates a contribution

$$\mathcal{D}_{\mathrm{relax},ij}(\hat{\rho}) = \begin{array}{c} \\ i \\ \\ \\ j \\ \\ \end{array}\!\!\begin{array}{c} \begin{array}{ccccccc} & & & i & & j & \\ \end{array} \\ \begin{bmatrix} & & & -\frac{1}{2}\gamma_{ij}\rho_{1j} & & \\ & & & \vdots & & \\ & & \gamma_{ij}\rho_{jj} & -\frac{1}{2}\gamma_{ij}\rho_{ij} & & \\ & & & \vdots & & \\ -\frac{1}{2}\gamma_{ij}\rho_{j1} & \cdots & -\frac{1}{2}\gamma_{ij}\rho_{ji} & \cdots & -\gamma_{ij}\rho_{jj} & \cdots & -\frac{1}{2}\gamma_{ij}\rho_{jN} \\ & & & \vdots & & & \\ & & & -\frac{1}{2}\gamma_{ij}\rho_{Nj} & & & \end{bmatrix} \end{array}, \tag{3.8}$$

where the mapping between the indices i, j and k mentioned before has been reverted and the scattering rates γ_{ij} have been introduced. These rates correspond to the coefficients c_{kk} for $k \geq N$ and represent the

relaxation (or scattering) processes from the basis state $|i\rangle$ to the basis state $|j\rangle$. Now we can perform the sum over all $\mathcal{D}_{\text{relax},ij}$ and find that the resulting expressions

$$\partial_t \rho_{ii}|_{\text{relax}} = \sum_{j=1, j\neq i}^{N} \gamma_{ji}\rho_{jj} - \tau_i^{-1}\rho_{ii} \tag{3.9}$$

for the time derivative of each population provide a clear and physically meaningful description of incoherent processes that is widely used in both theoretical as well as experimental work [9], [32]. Here, the inverse population life times $\tau_i^{-1} = \sum_{j=1, j\neq i}^{N} \gamma_{ij}$ have been introduced for brevity. We derive in similar fashion the time derivatives of the coherence terms $\rho_{ij,i\neq j}$, which read

$$\partial_t \rho_{ij,i\neq j}\big|_{\text{relax}} = -\frac{1}{2}\left(\tau_i^{-1} + \tau_j^{-1}\right)\rho_{ij}. \tag{3.10}$$

Therefore, the relaxation or scattering processes lead also to the loss of coherence, which is commonly referred to as dephasing.

As next step, we plug in the diagonal basis operators into Eq. (2.83). It has been pointed out that assuming a completely diagonal coefficient matrix C restricts the resulting description too strongly [57], [58]. Therefore, the resulting contribution to the dissipation superoperator reads

$$\mathcal{D}_{\text{deph}}(\hat{\rho}) = -\frac{1}{2}\sum_{m,n=1}^{N-1} c_{mn} \begin{bmatrix} 0 & \left(F_{m,11} - F_{n,22}\right)^2 \rho_{12} & \left(F_{m,11} - F_{n,33}\right)^2 \rho_{13} & \dots \\ \left(F_{m,22} - F_{n,11}\right)^2 \rho_{21} & 0 & \left(F_{m,22} - F_{n,33}\right)^2 \rho_{23} & \dots \\ \left(F_{m,33} - F_{n,11}\right)^2 \rho_{31} & \left(F_{m,33} - F_{n,22}\right)^2 \rho_{32} & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{3.11}$$

where the $F_{k,ii}$ denote the diagonal elements of the diagonal operators $\hat{F}_{k,k\leq N-1}$. We can readily see that the populations are not affected by this contribution, but the coherence terms experience additional dephasing

$$\partial_t \rho_{ij,i\neq j}\big|_{\text{deph}} = -\frac{1}{2}\sum_{m,n=1}^{N-1} c_{mn}\left(F_{m,ii} - F_{n,jj}\right)^2 \rho_{ij} = -\gamma_{ij,\text{p}}\rho_{ij}, \tag{3.12}$$

which is usually referred to as pure dephasing with the pure dephasing rate $\gamma_{ij,\text{p}}$.

At this point, we have assembled a physically meaningful description with the help of the chosen basis operators, but the values of the scattering rates γ_{ij} and the pure dephasing rates $\gamma_{ij,\text{p}}$ remain to be determined. Fortunately, the description is compatible with all charge carrier simulation methods discussed above. For example, the expressions in Eq. (3.9) can be found in classical rate equation approaches (cf. [10]), where the rates are determined empirically. By choosing appropriate values for the pure dephasing rates in similar fashion, it is possible to extend these approaches to an empirical, yet fully quantum mechanical treatment. For self-consistent simulations, on the other hand, NEGF methods [53], density matrix approaches (e.g., [50]), or ensemble Monte-Carlo methods (see e.g., [54], [59]) can be employed to determine the scattering rates γ_{ij}. The pure dephasing rates can be calculated in self-consistent manner using Ando's model [60]–[62], in which the dephasing rates

$$\gamma_{ij,\text{deph}} = \gamma_{ji,\text{deph}} = \frac{1}{2}\left(\tau_i^{-1} + \tau_j^{-1}\right) + \gamma_{ij,\text{p}} \tag{3.13}$$

are compatible with Eqs. (3.10) and (3.12). For example, the density matrix EMC technique presented in [54] uses this approach.

Independent of how the scattering and pure dephasing rates have been determined, it remains to be verified whether the resulting master equation yields a completely positive and trace-preserving solution. We recall that the coefficient matrix C must be positive semi-definite (cf. Theorem 2.2.2) in order to fulfill this requirement. Since the coefficients related to the off-diagonal basis operators are only non-zero on the main diagonal of C, it is sufficient that all scattering rates $\gamma_{ij} \geq 0$ are real and non-negative. For the coefficients related to the diagonal basis operators the situation is more complicated [57], [58]. First, the

pure dephasing rates must be converted into a $(N-1)\times(N-1)$ block matrix with the entries c_{mn}. Since there are $(N-1)^2$ coefficients and $N(N-1)/2$ pure dephasing rates, the problem is overdetermined and it is not guaranteed that the conversion succeeds. After a successful conversion, the eigenvalues of the block matrix can be determined and inspected [57]. If they are non-negative, the complete coefficient matrix is positive semi-definite and the values for the pure dephasing rates constitute a reasonable choice.

Reduction to a single representative system

We note that the density matrix in the description above could easily contain hundreds of relevant energy levels and causes a significant computational workload. Also, the boundary conditions of the description need to be discussed. Due to the periodicity of the heterostructure, however, it is reasonable to assume that the energy levels of a single period represent the complete system. This goes hand in hand with the usually applied periodic boundary conditions [9]. Then, the description is reduced to a $N'\times N'$ density matrix of the representative system, where N' is the number of levels per period (usually in the single-digit range). The coherence terms related to levels of different periods are neglected as a consequence. For the remaining elements, $\rho_{ij}=\rho_{ij}^{-}=\rho_{ij}^{+}$ is assumed, where the $\rho_{ij}^{\pm}$ denote the the density matrix entries related to the adjacent periods.

The dynamical behavior of the representative system is still governed by the master equation (3.5) with the following considerations in mind. In the majority of related work a block-diagonal Hamiltonian $\hat{H}$ is implicitly assumed, where the identical blocks correspond to the Hamiltonian $\hat{H}'$ of the representative system. Since both the density matrix and the Hamiltonian of the complete system are assumed to be block-diagonal, the master equation

$$\partial_t\hat{\rho}' = -\mathrm{i}\hbar^{-1}\left[\hat{H}',\hat{\rho}'\right] + \mathcal{D}'(\hat{\rho}) \tag{3.14}$$

can be extracted from Eq. (3.5), where the dissipation superoperator $\mathcal{D}'$ needs further consideration. As already pointed out, this assumption neglects the coherent interaction between different periods. Since the period boundaries can be chosen arbitrarily, it is usually possible to choose the periods such that this assumpion is valid. The incoherent interaction between the periods, on the other hand, cannot be ignored, but is often assumed to be limited to the nearest neighbors [9]. For this case, Eq. (3.9) reads

$$\partial_t\rho_{ii}|_{\mathrm{relax}} = \sum_{j=1,j\neq i}^{N'}\left(\gamma_{ji}\rho_{jj}-\gamma_{ij}\rho_{ii}\right) + \sum_{j=1}^{N'}\left(\gamma_{ji}^{+}\rho_{jj}^{-}-\gamma_{ij}^{-}\rho_{ii}\right) + \sum_{j=1}^{N'}\left(\gamma_{ji}^{-}\rho_{jj}^{+}-\gamma_{ij}^{+}\rho_{ii}\right), \tag{3.15}$$

where the rates $\gamma_{ji}^{\pm}$ represent the scattering from a period to the next and to the previous period, respectively. We use our assumption $\rho_{ij}=\rho_{ij}^{-}=\rho_{ij}^{+}$ and simplify the equation to

$$\partial_t\rho_{ii}|_{\mathrm{relax}} = \sum_{j=1,j\neq i}^{N'}\underbrace{\left(\gamma_{ji}+\gamma_{ji}^{+}+\gamma_{ji}^{-}\right)}_{\gamma'_{ji}}\rho_{jj} - \underbrace{\sum_{j=1,j\neq i}^{N'}\left(\gamma_{ij}+\gamma_{ij}^{+}+\gamma_{ij}^{-}\right)}_{\tau_i'^{-1}}\rho_{ii}, \tag{3.16}$$

which still is in the form of Eq. (3.9) with different parameters γ'_{ji} and $\tau_i'^{-1}$. The same holds in principle for Eq. (3.10), which now reads

$$\partial_t\rho_{ij,i\neq j}\big|_{\mathrm{relax}} = -\frac{1}{2}\left(\tau_i'^{-1}+\tau_j'^{-1}+\gamma_{ii}^{-}+\gamma_{ii}^{+}+\gamma_{jj}^{-}+\gamma_{jj}^{+}\right)\rho_{ij}. \tag{3.17}$$

Here, it makes sense to incorporate the additional terms into the pure dephasing rate

$$\gamma'_{ij,\mathrm{p}} = \gamma'_{ij,\mathrm{p}} + \frac{1}{2}\left(\gamma_{ii}^{-}+\gamma_{ii}^{+}+\gamma_{jj}^{-}+\gamma_{jj}^{+}\right) \tag{3.18}$$

in order to maintain the form of the description. We note that existing models in related literature often implicitly apply further approximations, in which certain additional scattering rates are neglected (e.g., in [23]).

3.1.3 Impact of the electromagnetic field on the time evolution

Now we introduce the effect of the electromagnetic field on the quantum mechanical system into the master equation (3.5). To this end, we recall the Hamiltonian in Eq. (2.48) and remember that a canonical transformation is required for the mechanical momentum $\boldsymbol{p}$ and canonical momentum $\boldsymbol{p'}$ to coincide (cf. Section 2.2.3). We can readily see from the transformation laws that solving the equation $\nabla\xi = -q\boldsymbol{A}$ will yield the required additional term for the Hamiltonian. At this point, the dipole approximation is usually invoked [32], which assumes that the vector potential field $\boldsymbol{A}$ does not vary significantly on the scale of the quantum mechanical system. This is on a par with the separation ansatz in Eq. (3.2), as we implicitly assume that the quantum mechanical system does not extend in-plane. Additionally, the field is assumed to be slowly varying in growth direction. Then, the field $\boldsymbol{A} \neq f(\boldsymbol{r})$ can be assumed constant in space and for the auxiliary function $\xi = -q\boldsymbol{r}\boldsymbol{A}$ holds. According to Eq. (2.53) the transformed Hamiltonian reads

$$\hat{H} = \frac{\hat{\boldsymbol{p}}^2}{2m} + q\varphi + V + q\hat{\boldsymbol{r}}\partial_t\boldsymbol{A}. \tag{3.19}$$

We plug in the electron charge for $q = -e$ and compare the Hamiltonian $\hat{H}$ with Eq. (3.4). Since the contribution of the scalar potential φ is already included in $\hat{H}_0$, we find that $V = V_\mathrm{c}$ and write for the total Hamiltonian

$$\hat{H} = \hat{H}_0 - \hat{\boldsymbol{\mu}}\boldsymbol{E}_\mathrm{O}, \tag{3.20}$$

where we introduce the dipole moment operator $\hat{\boldsymbol{\mu}} = -e\hat{\boldsymbol{r}}$ and use $\boldsymbol{E}_\mathrm{O} = -\partial_t\boldsymbol{A}$, as described by Eq. (3.1).

As already mentioned above, we assume that the optical field does not affect the basis states and can be seen as time-dependent perturbation. Then, the ansatz in Eq. (3.2) can be used to write the dipole moment operator as matrix with the elements

$$\hat{\boldsymbol{\mu}}_{ij} = -e\langle i|\boldsymbol{r}|j\rangle = -eS^{-1}\int \psi_i^*(z)\boldsymbol{r}\psi_j(z)\mathrm{d}\boldsymbol{r}. \tag{3.21}$$

Here, we assume that the interaction with the electromagnetic field preserves the in-plane wave numbers k_x and k_y and drop the subscript $\boldsymbol{k}$ from the wave function envelopes. Since the envelopes only depend on the growth direction, the in-plane components of the dipole operator

$$\hat{\mu}_{ij,p} = -eS^{-1}\underbrace{\int p\mathrm{d}x\mathrm{d}y}_{=0}\int \psi_i^*(z)\psi_j(z)\mathrm{d}z, \tag{3.22}$$

where $p = \{x, y\}$, vanish. The component in growth direction, on the other hand, can be expressed as

$$\hat{\mu}_{ij,z} = -eS^{-1}\int \mathrm{d}x\mathrm{d}y\int \psi_i^*(z)z\psi_j(z)\mathrm{d}z = -e\int \psi_i^*(z)z\psi_j(z)\mathrm{d}z, \tag{3.23}$$

where the normalization condition was applied to simplify the result. As a consequence, the Hamiltonian can be reduced to

$$\hat{H} = \hat{H}_0 + \hat{\mu}_z\partial_t A_z = \hat{H}_0 - \hat{\mu}_z E_{\mathrm{O},z} \tag{3.24}$$

in this basis.

As a side note, it should be mentioned that some methods consider the optical field during the determination of the dissipation superoperator [9]. In particular, this is relevant for quantum cascade lasers that feature large output power, which is (at the current state of research) true for mid-infrared QCLs.

3.1.4 Impact of the quantum mechanical systems on the electromagnetic field

Conversely, the quantum mechanical systems affect the electromagnetic field by means of the source terms in Maxwell's equations (2.2), namely the charge density

$$\rho = -\nabla\boldsymbol{P} + \rho_\mathrm{f} \tag{3.25}$$

and the current density

$$\boldsymbol{J} = \nabla \times \boldsymbol{M} + \partial_t \boldsymbol{P} + \boldsymbol{J}_\mathrm{f}, \tag{3.26}$$

where the definitions from Section 2.1.1 have been used. First of all, the magnetization $\boldsymbol{M}$ can be safely neglected for optical frequencies [9]. Then, the question is raised whether the charge carriers described by the quantum mechanical systems should be treated as bound or free carriers. However, a clear distinction is neither possible in the scope of modeling quantum cascade lasers, nor is it necessary, since the contributions of bound and free carriers always appear as sum [32]. Therefore, we can safely assume that the charge carriers are (quasi-)bound and proceed with the description of the charge density $\rho_\mathrm{qm} = -\nabla \boldsymbol{P}_\mathrm{qm}$ and the corresponding polarization density $\boldsymbol{P}_\mathrm{qm}$. We note that in this case the continuity equation (2.4) is fulfilled by definition.

The charge carriers that are described by the quantum mechanical systems have been introduced by doping. Therefore, we need to take into account the distribution $n_\mathrm{D}(z)$ of positive donors, which is static and varies only in growth direction z. The distribution of their negative counterpart, namely the electrons, can be described using the quantum mechanical probability distribution $|\Psi_i(\boldsymbol{r})|^2$ (as postulated in Section 2.2.1) of electrons in the state $|\Psi_i\rangle$. Here, we assume that the total number of electrons equals to the total number of donors

$$N_\mathrm{e} = N_\mathrm{D} = \int_V n_\mathrm{D}(z) \mathrm{d}\boldsymbol{r} \tag{3.27}$$

where V is the volume of the active region. Furthermore, each basis state $|\Psi_i\rangle$ is populated by $N_\mathrm{e} p_i$ electrons, where we use the population probability values p_i with $\sum_i p_i = 1$. Using the ansatz in Eq. (3.2) and dropping the subscript $\boldsymbol{k}$, we can determine the contribution

$$\rho_\mathrm{qm,e} = -e \sum_i N_\mathrm{e} p_i \, |\Psi_i(\boldsymbol{r})|^2 = -e \sum_i N_\mathrm{D} p_i S^{-1} \, |\psi_i(z)|^2 \tag{3.28}$$

of the electrons to the overall charge distribution

$$\rho_\mathrm{qm}(z) = e \left[n_\mathrm{D}(z) - \sum_i n_i^\mathrm{S} \, |\psi_i(z)|^2 \right], \tag{3.29}$$

where the sheet densities $n_i^\mathrm{S} = p_i N_D S^{-1}$ are introduced to match the expressions in related literature [9], [63], [64]. The sheet density values (or, essentially, the probabilities p_i) are generally determined by charge carrier simulations, which usually calculate the stationary distribution [9]. It should be pointed out, however, that the probabilities may be time-dependent. In order to acquire an initial estimate of the probability values, Fermi-Dirac statistics are frequently applied [63], [64].

Since the positive and negative charges are distributed differently, the overall charge distribution ρ_qm causes a polarization component

$$\begin{aligned} \boldsymbol{P}_\mathrm{qm} &= \lim_{\Delta V \to 0} \frac{1}{\Delta V} \int_{\Delta V} \rho_\mathrm{qm}(z) \boldsymbol{r} \mathrm{d}\boldsymbol{r} \\ &= \lim_{\Delta V \to 0} \frac{1}{\Delta V} e \left[\int_{\Delta V} n_\mathrm{D}(z) \boldsymbol{r} \mathrm{d}\boldsymbol{r} - \int_{\Delta V} n_\mathrm{D}(z) \mathrm{d}\boldsymbol{r} \sum_i p_i \int_{\Delta V} \Psi_i^*(\boldsymbol{r}) \boldsymbol{r} \Psi_i(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \right] \end{aligned} \tag{3.30}$$

according to Eq. (2.14), where the number of electrons N_e has been adjusted to the number of donors in the test volume ΔV. At this point it makes sense to discuss this test volume, which defines the resolution of the polarization from a macroscopic perspective. From the microscopic point of view, on the other hand, it is crucial that the test volume is large enough to contain the complete domain of the quantum mechanical probability distribution function. Frequently, the period length L_p is used as a compromise between the two perspectives, as the charge distribution is assumed to be periodic with the length L_p [9]. With a suitable test volume, the limit operation has been applied successfully and can be dropped. On the other hand, the boundaries of the integral over the quantum mechanical probability distribution can be ignored. Then, the contribution of the electrons to the polarization

$$\boldsymbol{P}_\mathrm{qm,e} = -\frac{1}{\Delta V} \int_{\Delta V} n_\mathrm{D}(z) \mathrm{d}\boldsymbol{r} \sum_i p_i \int \Psi_i^*(\boldsymbol{r}) e \boldsymbol{r} \Psi_i(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} = n_\mathrm{3D} \, \mathrm{Tr}\left\{ \hat{\boldsymbol{\mu}} \hat{\rho} \right\} \tag{3.31}$$

can be related to Eq. (2.59), which defines the macroscopic value of a quantum mechanical quantity as product of the particle density

$$n_{3D} = \frac{1}{\Delta V}\int_{\Delta V} n_D(z)\mathrm{d}\boldsymbol{r} \approx \frac{1}{L_p}\int_{L_p} n_D(z)\mathrm{d}z \tag{3.32}$$

and its expectation value (in this case, that of the dipole moment operator $\hat{\boldsymbol{\mu}} = -e\hat{\boldsymbol{r}}$)

$$\langle\hat{\boldsymbol{\mu}}\rangle = \mathrm{Tr}\left\{\hat{\boldsymbol{\mu}}\hat{\rho}\right\} = -e\langle\hat{\boldsymbol{r}}\rangle = -e\sum_i p_i \langle\Psi_i|\hat{\boldsymbol{r}}|\Psi_i\rangle, \tag{3.33}$$

as defined in Eq. (2.62). The exact form of the contribution of the donors is not derived since only the time derivative of the polarization enters Eq. (3.26) in the form of the polarization current

$$\partial_t \boldsymbol{P}_{qm} = n_{3D}\,\mathrm{Tr}\left\{\hat{\boldsymbol{\mu}}\partial_t\hat{\rho}\right\}. \tag{3.34}$$

As the charge distribution only depends on the growth direction, simplifications similar to the dipole moment in Eq. (3.22) can be made. We expand Eq. (3.30) componentwise for the in-plane directions $p = \{x, y\}$

$$P_{qm,p} = \lim_{\Delta V\to 0}\frac{1}{\Delta V}\underbrace{\int p\mathrm{d}x\mathrm{d}y}_{=0}\int \rho_{qm}(z)\mathrm{d}z = 0 \tag{3.35}$$

and can readily see that the corresponding components vanish. The remaining polarization component in growth direction leads to a polarization current component

$$\partial_t P_{qm,z} = n_{3D}\,\mathrm{Tr}\left\{\hat{\mu}_z\partial_t\hat{\rho}\right\}, \tag{3.36}$$

where only the component $\hat{\mu}_z$ of the dipole operator in growth direction is considered.

With the contributions ρ_{qm} and $\partial_t P_{qm,z}$ of the quantum mechanical systems to the charge density and current density we proceed to the next section, where those contributions are included in the model of the electromagnetic fields in the resonator.

3.2 Electromagnetic field in the resonator cavity

In the optical resonator the electromagnetic field consists of different components. As already pointed out in Eq. (3.1), there are the components $\boldsymbol{E}_O$ and $\boldsymbol{B}_O$ that represent the optical field in the terahertz or mid-infrared regime. Additionally, one component should be considered that stands for the bias voltage, which is applied to the quantum cascade laser. The bias voltage is assumed to be constant in time and is therefore likely to induce an electric field $\boldsymbol{E}_B$, but no magnetic field. With Eq. (2.19) in mind, we find that this can be reflected by the definitions

$$\boldsymbol{E}_B = -\nabla\varphi, \tag{3.37a}$$

$$\boldsymbol{E}_O = -\partial_t\boldsymbol{A}, \tag{3.37b}$$

$$\boldsymbol{B}_O = \nabla\times\boldsymbol{A}. \tag{3.37c}$$

Here, we assume a time-independent scalar potential $\varphi(z)$ that only varies in growth direction z. This is a reasonable assumption since the bias voltage is applied to contacts with high conductance (usually, gold or copper) and does not vary in time. Then, the sum $\boldsymbol{E}_B + \boldsymbol{E}_O = \boldsymbol{E}$ yields the total electric field, and the total magnetic field $\boldsymbol{B} = \boldsymbol{B}_O$ equals the optical magnetic field. As a consequence, we can divide the discussion into two parts, namely the description of the bias field and the propagation of the optical field.

In this discussion, we include the microscopic contributions of the quantum mechanical systems to charge density and current density presented in the previous section, as well as macroscopic contributions that stem from the bulk properties of the resonator materials. Since the magnetization can be safely neglected, it is sufficient to consider a background electric susceptibility χ_e and allow for ohmic losses σ. It should be pointed out, however, that these two properties may be frequency dependent and may vary in space, although they can usually be assumed constant in propagation direction x.

3.2.1 Description of the bias field

In Coulomb gauge, the scalar potential can be described by the Poisson equation

$$-\nabla^2\varphi = \varepsilon_0^{-1}\rho, \tag{3.38}$$

which corresponds to Eq. (2.31b) after dropping the prime. Here, the charge density

$$\rho = -\nabla \boldsymbol{P} = -\nabla \boldsymbol{P}_{\mathrm{bg}} + \rho_{\mathrm{qm}} \tag{3.39}$$

contains the contribution of the quantum mechanical systems as well as the the isotropic background polarization $\boldsymbol{P}_{\mathrm{bg}} = \varepsilon_0\chi_{\mathrm{e}}(z)\boldsymbol{E}_{\mathrm{B}}$, which is related to the static bias field in this case. Therefore, the electric susceptibility χ_{e} is the static electric susceptibility and varies in growth direction z due to the different layers of the heterostructure. Using these definitions, we can write Eq. (3.38) as

$$-\varepsilon_0\nabla^2\varphi = \varepsilon_0\nabla\chi_{\mathrm{e}}(z)\nabla\varphi + \rho_{\mathrm{qm}}, \tag{3.40}$$

and, with $\nabla^2\varphi + \nabla\chi_{\mathrm{e}}(z)\nabla\varphi = \nabla\varepsilon_{\mathrm{r}}(z)\nabla\varphi$ and Eq. (3.29), as

$$-\varepsilon_0\nabla\varepsilon_{\mathrm{r}}(z)\nabla\varphi = e\left[n_{\mathrm{D}}(z) - \sum_i n_i^{\mathrm{S}}\,|\psi_i(z)|^2\right]. \tag{3.41}$$

Since the bias field is assumed to have only a non-zero component in growth direction z, the scalar potential φ does not depend on x and y (cf. Eq. (3.37)). Therefore, Eq. (3.41) can be brought into the one-dimensional form

$$-\varepsilon_0\partial_z\varepsilon_{\mathrm{r}}(z)\partial_z\varphi = e\left[n_{\mathrm{D}}(z) - \sum_i n_i^{\mathrm{S}}\,|\psi_i(z)|^2\right] \tag{3.42}$$

and determines the scalar potential together with the boundary conditions $\varphi(0) = 0$ and $\varphi(L_z) = U_{\mathrm{B}}$.

3.2.2 Propagation of the optical field

As defined in Eq. (3.37), the optical part of the electromagnetic field is represented by the vector potential $\boldsymbol{A}$. The vector potential is governed by

$$\nabla^2\boldsymbol{A} - \varepsilon_0\mu_0\partial_t^2\boldsymbol{A} - \varepsilon_0\mu_0\partial_t\nabla\varphi = -\mu_0\boldsymbol{J}, \tag{3.43}$$

i.e., the second part of the potential formulation of Maxwell's equations in Coulomb gauge, cf. Eq. (2.31b). Since the bias field is assumed to be static, $\partial_t\nabla\varphi = 0$ holds and the equation can be simplified to

$$\nabla^2\boldsymbol{A} - \varepsilon_0\mu_0\partial_t^2\boldsymbol{A} = -\mu_0\boldsymbol{J}, \tag{3.44}$$

a wave equation with the source term $-\mu_0\boldsymbol{J}$. At this point, we assume that the magnetization is negligible at optical frequencies [9], i.e., $\boldsymbol{B} = \mu_0\boldsymbol{H}$ or $\boldsymbol{M} = 0$. Hence, the bound current density is $\boldsymbol{J}_{\mathrm{b}} = \partial_t\boldsymbol{P}$, where we assume that the polarization $\boldsymbol{P} = \varepsilon_0\chi_{\mathrm{e}}\boldsymbol{E}_{\mathbf{O}} + \boldsymbol{P}_{\mathrm{qm}}$ again consists of an isotropic background polarization and the contribution of the quantum mechanical systems. Here, the electric susceptibility χ_{e} may differ from its static value in the section before. Additionally, we allow for ohmic losses by adding the free current density $\boldsymbol{J}_{\mathrm{f}} = \sigma\boldsymbol{E}_{\mathbf{O}}$. After plugging those relations into Eq. (3.44), we find that the wave equation now reads

$$\nabla^2\boldsymbol{A} - \varepsilon_0\mu_0\partial_t^2\boldsymbol{A} = \mu_0\left[\varepsilon_0\chi_{\mathrm{e,O}}\partial_t^2\boldsymbol{A} - \partial_t\boldsymbol{P}_{\mathrm{qm}} + \sigma\partial_t\boldsymbol{A}\right] \tag{3.45}$$

and can be simplified to

$$\nabla^2\boldsymbol{A} - \varepsilon_0\varepsilon_{\mathrm{r}}\mu_0\partial_t^2\boldsymbol{A} - \mu_0\sigma\partial_t\boldsymbol{A} = -\mu_0\partial_t\boldsymbol{P}_{\mathrm{qm}}. \tag{3.46}$$

The resulting wave equation for the vector potential can be transformed back to Maxwell's equations in macroscopic form and solved numerically. The three-dimensional approach allows a detailed model, but may be too demanding in terms of computational effort.

Determining the transversal modes

The typical wave guide geometry of quantum cascade lasers allows a significant simplification. The cross section can be assumed to be invariant in propagation direction, which means that the conductivity $\sigma(y,z)$, the susceptibility $\chi_{\mathrm{e,O}}(y,z)$, and, as a consequence, the permittivity $\varepsilon_{\mathrm{r}}(y,z)$ do not depend on x. Then, we can use the ansatz

$$A_p(\boldsymbol{r},t) = \Re\left\{\underline{A}_{\mathrm{T},p}(y,z)\underline{A}_{\mathrm{L},p}^{\underline{\beta},\omega}\exp\left(\mathrm{i}\underline{\beta}x - \mathrm{i}\omega t\right)\right\}, \tag{3.47}$$

for the components A_p of the vector potential in the directions $p = \{x,y,z\}$. This ansatz separates the vector potential into a static transversal distribution and a longitudinal, propagating component, with the complex amplitude $\underline{A}_{\mathrm{L},p}^{\underline{\beta},\omega}$ and the (generally complex) propagation constant $\underline{\beta}$. We assume that the source term $\partial_t \boldsymbol{P}_{\mathrm{qm}}$ in Eq. (3.46) is small and its effect on the transversal distributions can be neglected. Therefore, we ignore it at this point and take it into account at a later stage using perturbation theory. Then, by inserting the ansatz in Eq. (3.47) into Eq. (3.46), we can derive a complex differential equation

$$\nabla_{\mathrm{T}}^2\underline{A}_{\mathrm{T},p} - \underline{\beta}^2\underline{A}_{\mathrm{T},p} + \varepsilon_0\varepsilon_{\mathrm{r}}\mu_0\omega^2\underline{A}_{\mathrm{T},p} + \mathrm{i}\omega\mu_0\sigma\underline{A}_{\mathrm{T},p} = 0 \tag{3.48}$$

for each transversal distribution, where we used the transversal differential operator $\nabla_{\mathrm{T}}^2 = \partial_y^2 + \partial_z^2$. Alternatively, the differential equation can be brought into the standard Helmholtz form

$$\nabla_{\mathrm{T}}^2\underline{A}_{\mathrm{T},p} = \left[\underline{\beta}^2 - \varepsilon_0\underline{\varepsilon}_{\mathrm{r}}(y,z)\mu_0\omega^2\right]\underline{A}_{\mathrm{T},p}, \tag{3.49}$$

where the complex permittivity $\underline{\varepsilon}_{\mathrm{r}}$ has been used for compactness. For anything but the most simple cases, Eq. (3.49) must be solved numerically with the boundary conditions $\underline{A}_{\mathrm{T},p}(y,z\to\pm\infty) = 0$. This step will be discussed in Section 3.3.2 and may yield more than one possible solution families for the transversal distributions $\underline{A}_{\mathrm{T},p}$, which are referred to as transversal modes in the following. Usually, lasers are designed to operate in a single transversal mode, which is normally the fundamental mode.

Reduction to a one-dimensional model

In the following we use the description of the transversal modes to derive a one-dimensional propagation model, which would be the result of starting again from Eq. (3.46) and invoking the plane wave approximation [10], in which the spatial derivatives $\partial_y = \partial_z = 0$ are neglected. This model would assume the form of a wave equation

$$\partial_x^2\underline{\mathbf{A}}^{\omega} + \varepsilon_0\underline{\varepsilon}_{\mathrm{eff}}\mu_0\omega^2\underline{\mathbf{A}}^{\omega} = \mathrm{i}\omega\mu_0\underline{\mathbf{P}}_{\mathrm{qm}}^{\omega} \tag{3.50}$$

in time-harmonic form, where the effective complex permittivity $\underline{\varepsilon}_{\mathrm{eff}}$ is spatially invariant. Since we learned in Section 3.1 that only the component A_z in growth direction is relevant for the time evolution of the quantum mechanical systems (cf. Eq. (3.24)) and the polarization density is only non-zero in growth direction (cf. Eq. (3.36)), we can restrict our considerations to the component A_z, which is described by the wave equation

$$\partial_x^2\underline{A}_z^{\omega} + \varepsilon_0\underline{\varepsilon}_{\mathrm{eff}}\mu_0\omega^2\underline{A}_z^{\omega} = \mathrm{i}\omega\mu_0\underline{P}_{\mathrm{qm},z}^{\omega}. \tag{3.51}$$

In order to bring our model above in the one-dimensional form in Eq. (3.51), we assume that we have determined the (generally frequency dependent) transversal distribution $\underline{A}_{\mathrm{T},z}$ and the propagation constant $\underline{\beta}$, and insert the more general ansatz

$$A_z(\boldsymbol{r},t) = \Re\left\{\underline{A}_{\mathrm{T},z}(y,z)\underline{A}_{\mathrm{L},z}^{\omega}(x)\exp\left(-\mathrm{i}\omega t\right)\right\} \tag{3.52}$$

into Eq. (3.46). By considering the polarization density in time-harmonic form, the result can be written as

$$\left(\nabla_{\mathrm{T}}^2\underline{A}_{\mathrm{T},z}\right)\underline{A}_{\mathrm{L},z}^{\omega} + \underline{A}_{\mathrm{T},z}\partial_x^2\underline{A}_{\mathrm{L},z}^{\omega} + \varepsilon_0\underline{\varepsilon}_{\mathrm{r}}\mu_0\omega^2\underline{A}_{\mathrm{T},z}\underline{A}_{\mathrm{L},z}^{\omega} = \underline{A}_{\mathrm{T},z}\left(\underline{\beta}^2\underline{A}_{\mathrm{L},z}^{\omega} + \partial_x^2\underline{A}_{\mathrm{L},z}^{\omega}\right) = \mathrm{i}\omega\mu_0\underline{P}_{\mathrm{qm},z}^{\omega} \tag{3.53}$$

with the Helmholtz equation (3.49). The result already resembles the wave equation (3.51) with

$$\underline{\beta}^2(\omega) = \varepsilon_0\underline{\varepsilon}_{\mathrm{eff}}(\omega)\mu_0\omega^2, \tag{3.54}$$

but the source term $\mathrm{i}\omega\mu_0 P^{\omega}_{\mathrm{qm},z}/\underline{A}_{\mathrm{T},z}$ still depends on the transversal coordinates y and z. Hypothetically, we could assume that the polarization density features the same transversal distribution as the vector potential and use a variation of the ansatz in Eq. (3.52) for $P_{\mathrm{qm},z}$ as well, which would instantly yield the desired one-dimensional wave equation. Considering that outside of the active region the electromagnetic optical field (represented by the vector potential) may be non-zero but the polarization density $P_{\mathrm{qm},z}$ is clearly zero, it becomes soon apparent that this ansatz would be an oversimplification.

Instead, we need to account for the overlap of the transversal distributions of the vector potential and the polarization density, respectively. It seems reasonable to assume that the transversal distribution of the polarization density is quite similar to that of the vector potential and can be expressed as $\underline{P}_{\mathrm{T},z}(y,z) = \underline{A}_{\mathrm{T},z}(y,z)\underline{C}(y,z)$, where $\underline{C}(y,z)$ is an arbitrary function over the transversal plane. Then, we can plug in the ansatz

$$P_{\mathrm{qm},z}(\boldsymbol{r},t) = \Re\left\{\underline{A}_{\mathrm{T},z}(y,z)\underline{C}(y,z)\underline{P}^{\underline{\beta},\omega}_{\mathrm{L},z}\exp\left(\mathrm{i}\underline{\beta}x - \mathrm{i}\omega t\right)\right\} \tag{3.55}$$

for the polarization density, together with the ansatz for the vector potential in Eq. (3.47), into Eq. (3.46), which yields

$$\left[\nabla_{\mathrm{T}}^2 - \underline{\beta}^2 + \varepsilon_0\underline{\varepsilon}_{\mathrm{r}}\mu_0\omega^2\right]\underline{A}_{\mathrm{T},z} = \mathrm{i}\omega\mu_0\underline{A}_{\mathrm{T},z}\underline{C}\underline{P}^{\underline{\beta},\omega}_{\mathrm{L},z}\left[\underline{A}^{\underline{\beta},\omega}_{\mathrm{L},z}\right]^{-1} \tag{3.56}$$

or, after rearranging the terms,

$$\left\{\nabla_{\mathrm{T}}^2 + \varepsilon_0\underline{\varepsilon}_{\mathrm{r}}\mu_0\omega^2 - \mathrm{i}\omega\mu_0\underline{C}(y,z)\underline{P}^{\underline{\beta},\omega}_{\mathrm{L},z}\left[\underline{A}^{\underline{\beta},\omega}_{\mathrm{L},z}\right]^{-1}\right\}\underline{A}_{\mathrm{T},z} = \underline{\beta}^2\underline{A}_{\mathrm{T},z}. \tag{3.57}$$

We already assumed that the effect of the polarization density on the transversal distribution can be neglected. The effect on the propagation constant, however, is accounted for using time-independent perturbation theory, which is frequently used in quantum mechanics [65]. For the eigenvalue $\underline{\beta}^2$ we can derive the first-order correction

$$\Delta\underline{\beta}^2 = -\mathrm{i}\omega\mu_0\underline{P}^{\underline{\beta},\omega}_{\mathrm{L},z}\left[\underline{A}^{\underline{\beta},\omega}_{\mathrm{L},p}\right]^{-1}\frac{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\underline{C}(y,z)\mathrm{d}y\mathrm{d}z}{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\mathrm{d}y\mathrm{d}z}, \tag{3.58}$$

where $\underline{A}_{\mathrm{T},z}$ still denotes the solutions of the unperturbed differential equation (3.49), and the double integral in the denominator accounts for the fact that the solutions may not be normalized. It has been pointed out that the eigenvalue problem in Eq. (3.57) may be non-Hermitian and, strictly speaking, a biorthogonal set of basis functions must be used. However, it makes sense to use the correction in Eq. (3.58) as an approximation to the exact term [32].

Now we re-derive Eq. (3.53) but this time considering the ansatz

$$P_{\mathrm{qm},z}(\boldsymbol{r},t) = \Re\left\{\underline{A}_{\mathrm{T},z}(y,z)\underline{C}(y,z)\underline{P}^{\omega}_{\mathrm{L},z}(x)\exp\left(-\mathrm{i}\omega t\right)\right\} \tag{3.59}$$

for the polarization density, which yields

$$\partial_x^2\underline{A}^{\omega}_{\mathrm{L},z} + \underline{\beta}^2\underline{A}^{\omega}_{\mathrm{L},z} + \Delta\underline{\beta}^2\underline{A}^{\omega}_{\mathrm{L},z} = 0 \tag{3.60}$$

after factoring out the transversal distribution $\underline{A}_{\mathrm{T},z}$. We note that $\underline{A}^{\underline{\beta},\omega}_{\mathrm{L},z} = \underline{A}^{\omega}_{\mathrm{L},z}\exp(\mathrm{i}\underline{\beta}x)$ and $\underline{P}^{\underline{\beta},\omega}_{\mathrm{L},z} = \underline{P}^{\omega}_{\mathrm{L},z}\exp(\mathrm{i}\underline{\beta}x)$ and plug the expression for the correction $\Delta\underline{\beta}^2$ into Eq. (3.60), which provides a one-dimensional model

$$\partial_x^2\underline{A}^{\omega}_{\mathrm{L},z} + \underline{\beta}^2\underline{A}^{\omega}_{\mathrm{L},z} = \mathrm{i}\omega\mu_0\frac{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\underline{C}(y,z)\mathrm{d}y\mathrm{d}z}{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\mathrm{d}y\mathrm{d}z}\underline{P}^{\omega}_{\mathrm{L},z} \tag{3.61}$$

in time-harmonic form that does not depend on the transversal coordinates y and z anymore.

Frequently, it is assumed that the polarization density is zero outside of the active region and has the same transverse dependence as the vector potential in the active region. In this case, the expression

$$\frac{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\underline{C}(y,z)\mathrm{d}y\mathrm{d}z}{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\mathrm{d}y\mathrm{d}z} = \frac{\iint_{A_{\mathrm{qm}}}|\underline{A}_{\mathrm{T},z}|^2\mathrm{d}y\mathrm{d}z}{\iint_{-\infty}^{\infty}|\underline{A}_{\mathrm{T},z}|^2\mathrm{d}y\mathrm{d}z} = \Gamma \tag{3.62}$$

can be simplified using a suitable function $\underline{C}(y, z)$ and related to the overlap factor Γ [32], which is commonly used in non-linear optics [66]. Here, A_{qm} denotes the cross section of the active region.

As next step, we can transform the one-dimensional model back to the time domain using Eq. (3.54) and replace the effective complex permittivity with the effective real permittivity $\varepsilon_{\mathrm{eff}}$ and the effective conductivity σ_{eff}. Here, we keep in mind that those quantities and the overlap factor Γ still depend on frequency in general and employ the convolution theorem [14]. The model then assumes the form

$$\partial_x^2 A_{\mathrm{L},z} - \mu_0 \varepsilon_0 \varepsilon_{\mathrm{eff}}(t) * \partial_t^2 A_{\mathrm{L},z} - \mu_0 \sigma_{\mathrm{eff}}(t) * \partial_t A_{\mathrm{L},z} = -\mu_0 \Gamma(t) * \partial_t P_{\mathrm{L},z}, \tag{3.63}$$

where $\varepsilon_{\mathrm{eff}}(t)$, $\sigma_{\mathrm{eff}}(t)$ and $\Gamma(t)$ denote the impulse responses whose Fourier transforms correspond to the frequency dependent quantities mentioned before. We note that in the case of frequency independent quantites the convolution operations $*$ can be replaced by plain multiplications with the respective constant parameter. Similar wave equations can be derived for the longitudinal components $A_{\mathrm{L},x}$ and $A_{\mathrm{L},x}$, where the contribution of the polarization density $\boldsymbol{P}_{\mathrm{qm}}$ can be ignored.

3.3 Overview of the coupled model and simulation software

In this section, we combine the models for the electron dynamics in the gain medium and the propagation of the electromagnetic field in the resonator cavity. To this end, we revisit the components of the model discussed in the sections before (namely the model for the basis states of the electrons, the master equation including the dissipation superoperator, the Poisson equation for the bias field, the Helmholtz equation for the transversal modes, and the one-dimensional wave equation for the optical radiation field) and rearrange them. Here, we can benefit from the aforementioned reductions from three to one dimension and exploit the assumption that certain quantities do not vary strongly with respect to time. As a consequence, we can divide the coupled model into static prerequisite components and the dynamic part, allowing for an efficient treatment of the latter. As already mentioned, all model components require numerical simulation tools. Therefore, a non-exhaustive overview of existing simulation software [67] is provided along the description of the components in the following.

3.3.1 The basis states and the dissipation superoperator

In Section 3.1 we found that the description of the electron dynamics in the gain medium can be reduced to the master equation (3.5), where the basis states of the electrons only consider the growth direction, and the effects in in-plane direction are represented by the dissipation superoperator $\mathcal{D}$. Also, we assumed later that the interaction with the electromagnetic field does not affect the basis states and the dissipation superoperator, but can be treated as perturbation. Therefore, we can consider them as static prerequisites for the master equation, i.e., we can determine them once and simply use them in the dynamic model. As already mentioned, there are several ways to determine them, but all approaches discussed above are compatible with the master equation (3.5) and can be used interchangeably.

In order to present one exemplary approach, we use the Ben Daniel-Duke model and determine the basis states ψ_i by solving the Schrödinger equation 3.3, which can be written as

$$\left[-\frac{\hbar^2}{2}\partial_z \frac{1}{m^*(z)}\partial_z + V_{\mathrm{c}}(z) - e\varphi(z)\right]\psi_i(z) = \epsilon_i \psi_i(z) \tag{3.64}$$

with the Hamiltonian $\hat{H}_0$ described in Eq. (3.4). As already pointed out, there are different models for the scalar potential φ. In the most accurate case, the scalar potential is determined by the Poisson equation (3.42). The resulting Schrödinger-Poisson equation exhibits a circular dependency, as the charge distribution in the Poisson equation depends on the basis states, which is often handled by solving iteratively the Schrödinger equation (3.64) and the Poisson equation (3.42) until convergence is obtained [9]. Different numerical methods have been published that solve the Schrödinger-Poisson equation, including the transfer matrix method (e.g., in [68]), the finite difference scheme (e.g., in [69]), and the shooting method (e.g., in [70]).

Apart from various commercial solver tools, there are open source simulation software projects that are dedicated to solving the Schrödinger(-Poisson) equation, such as the aestimo project [70], [71].

The dissipation superoperator can be determined, for example, with the density matrix ensemble Monte Carlo (dmEMC) approach presented in [54]. Since it requires the basis states and determines (apart from the superoperator) the sheet densities n_i^{S} in Eq. (3.42), there is another circular dependency. This dependency is usually resolved by choosing suitable initial values for the sheet densities (often given by Fermi-Dirac statistics) and invoking the Schrödinger-Poisson solver and the dmEMC solver iteratively until the results converge.

3.3.2 Wave guide model

Similar to the separation of the in-plane effects from the electron dynamics in growth direction, the model for the electromagnetic field in the resonator cavity features a reduction in dimensionality. As discussed in Section 3.2, the typical wave guide geometry of the QCL resonator cavities leads to the formation of transversal modes that can be assumed to be constant with respect to time. Therefore, we can consider the transversal modes, which are the solution of the Helmholtz equation (3.49), as static prerequisite for the wave equation, which represents the dynamic part of the electromagnetic model. The complex eigenvalue problem in Eq. (3.49) appears quite often in nonlinear optics and is not necessarily restricted to QCLs. There exist semi-analytic approaches such as the film mode matching method [72] or the wave-matching method (WMM) [73], but it is also feasible to solve the problem using standard (commercial or open-source) finite-difference simulation tools. In terms of open-source simulation software projects, the WMM solver deserves mention at this point [74]. Finally, it should be mentioned that Eq. (3.49) can be simplified by neglecting the lateral direction, i.e., $\partial_y = 0$. The resulting one-dimensional equation models a slab wave guide and can be solved using transfer matrix methods [9].

3.3.3 The generalized Maxwell-Bloch equations

Basing on these static prerequisite components, we can combine the master equation (3.5) with the one-dimensional wave equation (3.63) to obtain the generalized Maxwell-Bloch equations. This combination constitutes the key model for the dynamical behavior of QCLs. We note that the Hamiltonian described in Eq. (3.24), which enters the master equation, requires the full vector potential A_z and depends in principle on $\boldsymbol{r}$. However, in the modeling of QCL dynamics usually one quantum mechanical system at $y = y_0, z = z_0$ represents the complete transversal plane. Due to performance reasons, the representative system is additionally constrained in the number of considered basis states (for example, to one representative period of the QCL gain medium) [32]. With this restriction, the master equation can be written as

$$\partial_t \hat{\rho} = -\mathrm{i}\hbar^{-1} \left[\left(\hat{H}_0 + \hat{\mu}_z \partial_t A_{\mathrm{L},z} \right), \hat{\rho} \right] + \mathcal{D}(\hat{\rho}), \tag{3.65}$$

where $\rho(x,t)$ and $A_{\mathrm{L},z}(x,t)$ depend on the propagation direction and the time, and $H_0(x)$ is time-independent. For the sake of completeness it should be noted that the contribution of the transversal distribution $A_{\mathrm{T},z}(y_0, z_0) = 1$ can be assumed to equal unity without loss of generality, as the linear Helmholtz equation allows for a multiplicative constant in the solutions. If the transversal distributions are used explicitly, care must be taken that this assumption is not violated.

As next step, we derive the term $\partial_t A_{\mathrm{L},z}$ by taking the time derivative of Eq. (3.63) and note that it corresponds to an electric field, which we denote by $-E_z$. This electric field is the solution of the resulting wave equation

$$\partial_x^2 E_z - \varepsilon_0 \varepsilon_{\mathrm{eff}}(t) * \mu_0 \partial_t^2 E_z - \mu_0 \sigma_{\mathrm{eff}}(t) * \partial_t E_z = \mu_0 \Gamma(t) * \partial_t^2 P_z, \tag{3.66}$$

where P_z denotes the longitudinal component of the polarization density that stems from the representative quantum mechanical system. This second-order partial differential equation can be rewritten (with the differentiation property of the convolution operation in mind) as system of two differential equations

$$\partial_x E_z = \mu_0 \partial_t H_y, \tag{3.67a}$$

$$\partial_x H_y = \varepsilon_0 \varepsilon_{\mathrm{eff}}(t) * \partial_t E_z + \sigma_{\mathrm{eff}}(t) * E_z + \Gamma(t) * \partial_t P_z, \tag{3.67b}$$

of first order that are, of course, Maxwell's equations in one-dimensional form. According to Eq. (3.36), the source term related to the polarization density is given by

$$\partial_t P_z = n_{3\mathrm{D}} \operatorname{Tr}\left\{\hat{\mu}_z \partial_t \hat{\rho}\right\}. \tag{3.68}$$

For the sake of completeness, we rewrite the master equation (3.65) as

$$\partial_t \hat{\rho} = -\mathrm{i}\hbar^{-1}\left[\hat{H}_0 - \hat{\mu}_z E_z, \hat{\rho}\right] + \mathcal{D}(\hat{\rho}). \tag{3.69}$$

Equations (3.67)-(3.69) are in the following referred to as the generalized Maxwell-Bloch equations in 1D. Here, the predicate "generalized" refers to the fact that the Maxwell-Bloch equations in their original form treated quantum mechanical systems with two basis states exclusively, whereas the form at hand is able to consider any number of basis states. The Maxwell-Bloch equations have been widely used to simulate the dynamical behavior of QCLs (e.g., in [9], [21]–[25], [75]–[80]) and can be considered as the standard model in the field. Therefore, they will be in the focus of the remainder of the thesis at hand and separate chapters shall be dedicated to the development of a solver tool and to the review and comparison of different numerical methods.

3.4 Summary

In this chapter, we have applied the theoretical basics to the problem of modeling the dynamical processes in a QCL. To this end, we have started with the basic assumption that the electromagnetic field consists of a static bias field and a dynamic radiation field. There are different approaches to model the dynamics of the electrons in the gain medium, i.e., the charge carrier transport in a biased semiconductor heterostructure. All of them share a common feature, namely the separation of the in-plane effects from the effects in growth direction. This separation allows a compact description of the electron dynamics with a reduced density matrix and a master equation of Lindblad type. To provide an exemplary approach to obtain the basis states of the electrons and the dissipation superoperator, which are the key elements in this master equation, we have discussed the Schrödinger equation in the scope of the Ben Daniel-Duke model in detail and established the link to the density matrix ensemble Monte Carlo (dmEMC) method. Here, it is assumed that only the static bias field affects the basis states and the dynamic radiation field can be treated as perturbation. Both fields and their interaction with the electron dynamics have been described in detail as next step. Here, the typical wave guide geometry of QCL resonator cavities has been exploited to separate the transversal distributions from the electromagnetic radiation field, resulting in an efficient, one-dimensional model.

During the discussion of the electron dynamics in the gain medium and the electromagnetic field in the resonator cavity, we have identified several components that form the complete model of the dynamical behavior of QCLs. Exploiting the reduction in dimensionality, we have extracted certain components, which can be considered as static prerequisites, from the dynamic model. Thereby, the dynamic model is reduced to the generalized Maxwell-Bloch equations in 1D, which have been identified as key model of the QCL dynamics. The components and the relations between them are depicted in Fig. 3.1, where the Schrödinger-Poisson model and the dmEMC method are used as example but could be replaced by alternatives if desired.

Finally, a brief overview of numerical methods and simulation software to solve the different model components numerically has been given. For the Maxwell-Bloch equations, which shall be in the focus of the remainder of the thesis at hand, separate chapters are dedicated to this discussion.

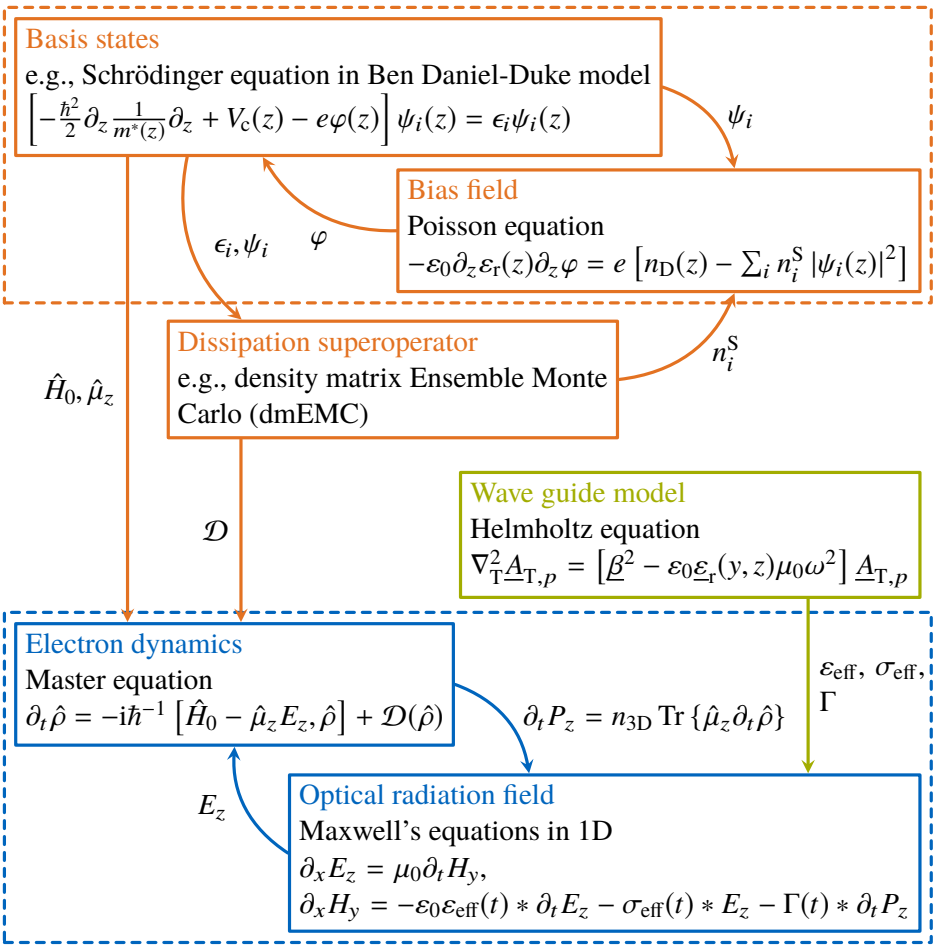


Figure 3.1 Exemplary mathematical model of the dynamical behavior of a quantum cascade laser. The model is divided into static prerequisite components and the dynamic part. The prerequisite components are the model for the basis states (e.g., the Schrödinger equation in the scope of the Ben Daniel-Duke model), which is often combined with the Poisson equation for the bias field (in this case resulting in the Schrödinger-Poisson model in orange dashed rectangle), the method that determines the dissipation superoperator $\mathcal{D}$ (e.g., the density matrix ensemble Monte Carlo method), and the Helmholtz equation for the transversal modes in the resonator cavity. The dynamic part (in the blue dashed rectangle) consists of the generalized Maxwell-Bloch equations in 1D. Adapted from M. Riesch, The QCL Stock Image Project [17], CC BY 4.0.

4 An open-source solver for the Maxwell-Bloch equations

> Freie Wissenschaft ist ebenso selbstverständlich wie freies Atmen.
>
> — *Lise Meitner*

> Talk is cheap. Show me the code.
>
> — *Linus Torvalds*

In the field of nonlinear optics, the Maxwell-Bloch equations are a valuable tool to model light-matter interaction [82], [83]. Originally devised to describe the behavior of magnetic moments of nuclei in a magnetic field [84], the Bloch equations soon found further application in which the dynamics of a two-level quantum mechanical system in resonance with an optical field are described [85]–[87]. This version is often referred to as optical Bloch equations and, coupled with Maxwell's equations for the optical field, was successfully applied to model nonlinear phenomena such as self-induced transparency [88], [89]. Later, a generalized form of the optical Bloch equations that considers an arbitrary number of energy levels was derived [90]. Instances of this form for three energy levels have been used e.g., to describe electromagnetically induced transparency and the related slow light propagation [91], [92]. Similarly, further application examples of this generalized form can be related to the propagation of light in different media, and considering quantum mechanical systems with two [93]–[95], three [96], [97] or six [98] energy levels, respectively.

While in earlier studies the media considered were mostly of gaseous form, advances in nanotechnology paved the way for solid state optoelectronic devices that exhibit (at least partially) coherent light-matter interaction, as is adequately described by the Maxwell-Bloch equations [32]. Here, the quantum cascade laser is a notable example. In the previous chapter, we identified the generalized Maxwell-Bloch equations as key model for the dynamical behavior of QCLs. As already pointed out, in many studies (e.g., [9], [21]–[25], [75]–[80]) this behavior was simulated in this framework. Also quantum dot devices have been extensively modeled based on Maxwell-Bloch equations [94], [99]–[101]. For a detailed overview of the applications, the reader is referred to a recent review paper [32] on the subject.

In the following, Section 4.1 reviews the one-dimensional generalized Maxwell-Bloch equations as they are used in the aforementioned related studies (and in particular, as they are used in the modeling of QCL dynamics). Since these equations cannot be solved analytically in anything but the most trivial cases, software tools are usually required to solver the equations numerically. Therefore, the requirements to the solver tools are discussed subsequently. As we shall see, the following overview of the state of the art in existing solver tools reveals that there is no solver tool that fulfills all requirements. As a consequence, Section 4.2 presents the implementation of mbsolve, an open-source solver for the Maxwell-Bloch equations, which was developed in the scope of the thesis at hand. In Section 4.3, the correct operation of the implementation is verified with the help of four application examples from related literature. At the same time, the various features of the mbsolve software are presented. Finally, after a short summary we conclude with an outlook on possible extensions of the software.

This chapter is largely based on M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations", *Comput. Phys. Commun.*, vol. 268, p. 108 097, Nov. 2021. DOI: 10.1016/j.cpc.2021.108097 in accordance with the Elsevier Copyright Policy.

4.1 The Maxwell-Bloch equations and existing solver tools

As already outlined, the Maxwell-Bloch equations describe the interaction of an electromagnetic field with quantum mechanical systems. In the most general picture, the electric field $\boldsymbol{E}(\boldsymbol{r},t)$ and the magnetic field $\boldsymbol{H}(\boldsymbol{r},t)$ depend on the three-dimensional space coordinate $\boldsymbol{r}$ and time t. The quantum mechanical systems are assumed to be uniformly distributed over space and are generally represented by the density matrix $\hat{\rho}(\boldsymbol{r},t)$. While the electromagnetic field is treated classically using the macroscopic Maxwell's equations (2.9) in 3D, the time evolution of the density matrix is governed by a quantum mechanical master equation of the form as given in Eq. (3.5). The Hamiltonian $\hat{H} = \hat{H}_0 - \hat{\boldsymbol{\mu}}\boldsymbol{E}$ in the master equation consists of the static Hamiltonian $\hat{H}_0$ and the interaction term $\hat{\boldsymbol{\mu}}\boldsymbol{E}$, as described in Section 3.1.3. At this point, the quantum mechanical systems are controlled by the electromagnetic field. The interaction cycle is completed by the introduction of a polarization term as given in Eq. (3.34) that enters Ampere's law in Maxwell's equations.

4.1.1 Generalized Maxwell-Bloch equations in 1D

The resulting three-dimensional form of the Maxwell-Bloch equations is quite complex. In the following we address a simplified version that is relevant for two classes of applications. The first class consists of simulation models in which the plane wave approximation [10] is a reasonable assumption. Those models can be found frequently in related literature (e.g., [93], [98], [102]). The second class of applications contains the simulation of optoelectronic devices with a wave guide geometry that permits the separation of transversal and longitudinal modes. As a consequence, only the propagation direction has to be considered in the Maxwell-Bloch equations, whereas the change in transversal directions is accounted for using effective electromagnetic properties (such as the effective permittivity ε_{eff} or the overlap factor Γ). In Chapter 3 we have found that the QCL is one example of such devices.

The simplified version of Maxwell-Bloch equations can be written in terms of the field components $E_z(x,t)$ and $H_y(x,t)$, where x is the propagation direction, and y, z are the transversal coordinates. Then, Maxwell's equations can be reduced to equations for the time evolution of the electric field

$$\partial_t E_z = \epsilon^{-1}\left(-\sigma E_z - \Gamma \partial_t P_z + \partial_x H_y\right) \tag{4.1}$$

and the magnetic field

$$\partial_t H_y = \mu^{-1}\partial_x E_z, \tag{4.2}$$

where the conductivity σ, the permittivity $\epsilon = \epsilon_0\epsilon_\text{r}$, and the permeability $\mu = \mu_0\mu_\text{r}$ can be related to a (generally complex) effective refractive index. Additionally, the overlap factor Γ accounts for a partial overlap of the transversal mode with the quantum mechanical systems and can be set to unity in cases where it is not required. The effective refractive index and the overlap factor are commonly used in optoelectronic device simulations and take into account the properties of the bulk material(s) in the setup as well as the wave guide geometry (if any). Generally, they depend on the position x and the frequency, which applies here for the electromagnetic properties ϵ, σ, μ and Γ. It should be noted that the frequency dependence reflects the chromatic dispersion of the background medium, which combines wave guide and bulk dispersion. We can readily see that Eqs. (4.1) and (4.2) correspond to Eqs. (3.67) if the frequency dependence is ignored. Indeed, we ignore this frequency dependence in the following for the sake of simplicity but note that the inclusion of the background dispersion should be in the focus of future work. The dispersion that stems from the quantum mechanical systems, on the other hand, is included in the polarization P_z and will be considered.

The variation of the material parameters in propagation direction is usually piecewise constant, in which case different regions of materials can be used to address this dependence. For each region, Eqs. (4.1) and (4.2) describe the electromagnetic field, where the material parameters are then constants. Within this model, the quantum mechanical systems are distributed only in the propagation direction and can be represented by the density matrix $\hat{\rho}(x,t)$. Furthermore, the dipole moment operator is now assumed to have only one non-zero component $\hat{\mu}_z$ [32]. As we have seen in Chapter 3, this is a reasonable assumption for QCLs. Therefore, the polarization contribution that stems from the quantum mechanical systems can

be determined with Eq. (3.68) and the time evolution of the density matrix is described by the master equation (3.69).

Now, as the system of partial differential equations is introduced, it is necessary to specify the initial and boundary conditions. The electric and magnetic fields require initial values $E(x, t = 0)$ and $H(x, t = 0)$, respectively. While many setups simply assume both to be zero, the fields may be initialized with random values to model spontaneous emission. The latter is typically applied during the simulation of lasers. As to boundary conditions, it is generally assumed that the electromagnetic wave is reflected with a certain reflectivity R at the simulation domain boundary. This assumption covers the case of perfect ($R = 1$) and semi-transparent ($0 < R < 1$) mirrors, which are often considered in optoelectronic devices, as well as perfectly matched layer (PML) boundary conditions ($R = 0$). It should be noted that the reflectivity values of the two boundaries $R_1 \neq R_2$ may be different. Since the master equation (3.69) does not include a spatial derivative, only the initial value of the density matrix at each point is required, such as thermal equilibrium (lowest energy level has the largest population) or inversion (some higher energy level, the so-called upper laser level, has the largest population). Finally, we note that source terms are often included in related literature. For example, an incoming electromagnetic pulse of Gaussian or sech shape is modeled by a source term in the electric field.

Together with those initial and boundary conditions, the system of partial differential equations form the generalized Maxwell-Bloch equations in 1D. If we reduce the model further by ignoring the propagation effects altogether, the equation system is nothing more than the master equation (3.69) that describes the time evolution for a single quantum mechanical system. As a consequence, the software project presented in the following at hand may also serve as solver for e.g., the Lindblad equation. Starting again from the one-dimensional version, we note that the generalized Maxwell-Bloch equations are able to treat an arbitrary number of energy levels N. The original Maxwell-Bloch equations can be derived by setting $N = 2$.

4.1.2 Requirements to the solver tool and existing approaches

At this point, the rotating wave approximation (RWA) in combination with the slowly varying envelope approximation (SVEA) is invoked in the majority of related studies to reduce further the complexity of the equation system. Naturally, these approximations may omit certain features of the solution [93]. Since it can be assumed that those features are crucial in the scope of simulations of quantum cascade laser frequency combs [103], [104], however we aim to avoid the RWA and SVEA in this work.

It should be noted that even when those approximations are used, analytical solutions are not generally available. As a consequence, we need to resort to numerical methods and solver tools that implement them. Although several numerical approaches for the Maxwell-Bloch equations have been discussed and compared in related literature [32], [93], [97], [98], [105]–[111], there is no definitive statement on what the best approach is. Also, just as with many problems in science and engineering, it is improbable that a single numerical method can cover all use cases. For example, while invoking the RWA/SVEA may not be suitable for quantum cascade laser frequency combs, it is perfectly reasonable for a multitude of problems. Therefore, the solver tool should provide support for multiple numerical methods in order to evaluate and compare different approaches. Here, a small and flexible code base is beneficial for rapid prototyping.

At the same time, the tool should permit productive usage with established numerical methods in the scope of our research, which leads to the following requirements. As already stated, a full-wave treatment of the optical field is desired (i.e., the rotating wave approximation should not be invoked). Secondly, as the full-wave treatment is computationally more intensive, the resulting numerical operations should be efficiently executed in parallel. Then, the solver should be able to deal with multiple sections of different materials, flexible initial and boundary conditions, source terms, and an arbitrary number of energy levels. Finally, the source code of the solver should be publicly available in order to allow extensions of the software and to improve the reproducibility of simulation results.

Unfortunately, the majority of the solver tools used in related work are not publicly available (e.g., in [93], [97], [98], [105]–[107]). In the following, the few exceptions to that rule should be mentioned. For example, the Electromagnetic Template Library (EMTL) [112] is a free C++ library with Message Passing Interface (MPI) support and has been used e.g., to model single quantum emitters [108]. However, it is

only available in binary form which makes it impossible to extend the library and port it to new computing architectures. The Freetwm [113] project is an open-source MATLAB code that solves the 1D Maxwell-Bloch equations. But to the best of our knowledge, it uses the rotating wave approximation and does not support alternative numerical methods. Also, solvers written and executed in MATLAB typically feature inferior performance than implementations in compiled languages such as C++. Finally, the open-source project MEEP [114] is a fully versed finite-difference time-domain (FDTD) solver for Maxwell's equations with parallelization support (using MPI) and a flexible user interface. Also, it features support for multi-level quantum mechanical systems. It is widely accepted in the simulation community and, therefore, it is a promising project to base future simulations onto. But one has to acknowledge that the evaluation and comparison of different numerical methods for the Maxwell-Bloch equations is hardly possible with such a large code base. This overview of existing approaches demonstrates the lack of a solver tool that fulfills all requirements, which lead to the development of a new tool named mbsolve.

4.2 Implementation of mbsolve

In this section, the architecture and the implementation details of the mbsolve software are presented. The modular architecture was designed with the requirements from Section 4.1 in mind, which we revisit briefly in the following. As already stated, it is crucial that mbsolve supports different numerical methods for the Maxwell-Bloch equations. We note that a flexible way to combine methods for Maxwell's equations and methods for the master equation would be beneficial. Then, for example, it would not be required to implement a certain method for Maxwell's equations multiple times when evaluating different algorithms to solve the master equation. Similarly, there are different parallelization techniques (OpenMP for shared memory systems, MPI for distributed memory systems, CUDA for NVIDIA graphics processing units (GPU), etc.), and mbsolve should be able to handle different techniques. Thereby, available simulation hardware can be targeted and exploited. A preliminary study found that the Maxwell-Bloch equations can be efficiently solved on GPUs [115], but the software should also work on a regular desktop PC without a high-end graphics card. As already mentioned, a small, yet flexible and extensible, code base is beneficial in order to achieve this. Finally, it is envisaged to share the resulting software project with the scientific community. Here, several measures must be taken to guarantee that other researchers can acquire, install and use the software [116]. While we discuss those measures in more detail below, the fundamental decision is the choice of programming language. The mbsolve software is written in the C++ programming language for performance reasons, but bindings for Python are offered in order to provide an easy-to-use interface for the researchers. Both programming languages are established in the scientific community and should constitute a reasonable choice [116].

During the implementation of the mbsolve software, the required flexibility was guaranteed by modularization and clearly defined interfaces. Those design criteria lead to the architecture depicted in Fig. 4.1. The mbsolve-lib base library constitutes the fundamental part of the software, as it provides an object oriented framework to define a simulation setup as well as the infrastructure to add solver and writer components. As the name suggests, the solver components implement numerical methods that solve the specified simulation setup using different parallelization techniques. After the solver has completed its work, the writer component is responsible for writing the simulation results into a file. In principle, writers for any file format can be implemented. However, open formats such as the Hierarchical Data Format (HDF) are clearly recommended [116].

In the following, the mbsolve-lib base library and a writer for the HDF file format are presented in more detail. Then, the measures taken to create a sustainable open-source software project out of mbsolve are discussed. Finally, installation notes and the requirements with respect to compilers and third-party software libraries are given.

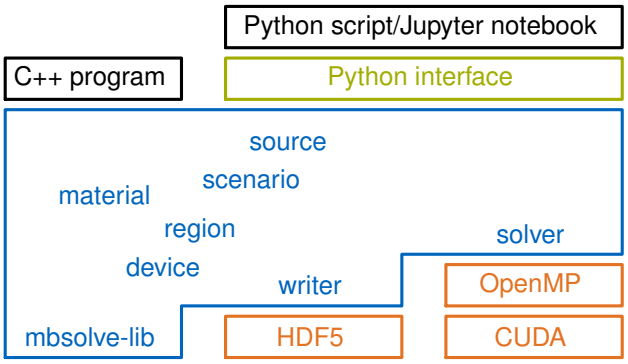


Figure 4.1 Overview of the mbsolve project. Reprinted from M. Riesch, The QCL Stock Image Project [17] (CC BY 4.0).

4.2.1 The mbsolve-lib base library

The object oriented framework that describes a simulation setup can be divided into two parts. The device part contains all properties of a setup that are static. These properties include, for example, the composition of the device under simulation in terms of regions and materials. The dynamic properties, on the other hand, are grouped into a scenario. For instance, simulation properties, such as the number of grid points used, or source terms are defined in the scenario. Thereby, it is possible to simulate the same device under different conditions, which was used, e.g., during the investigation of seeding effects in a quantum cascade laser [24]. After device and scenario are defined, the user can pass them to a solver, which calculates the simulation results, and optionally to a writer, which exports the results to a file.

Device setup and boundary conditions

The device is represented by a class of the same name. It contains a collection of **region** objects, which models a section of the device in which the material parameters are constant. A region is defined by the x_{start} and x_{end} coordinates as well as a pointer to a certain material. The envisaged materials have to be present in a static collection when the region is created. Thereby, material parameters are not directly stored in the regions. This enables the efficient treatment of periodic structures, where only few materials are repeatedly used in many different regions.

Each material is an instance of the eponymous class, which contains the electromagnetic properties as well as the description of the quantum mechanical systems. The former consist of the permittivity ϵ, the permeability μ, the overlap factor Γ, and the linear loss term

$$\alpha_0 = \frac{\sigma}{2\epsilon c} = \sqrt{\frac{\mu}{\epsilon}}\frac{\sigma}{2}, \tag{4.3}$$

which is related to the conductivity σ [10]. Here, $c = (\mu\epsilon)^{-1/2}$ is the speed of light in the material. The latter is incorporated as pointer to the class **qm_description**, which represents a quantum mechanical system. Here, the usage of a pointer allows polymorphism, i.e., the quantum mechanical description can assume different forms.

In the general form, this description requires the density of quantum mechanical systems n_{3D}, the Hamiltonian $\hat{H}_0$ and the dipole moment operator $\hat{\mu}_z$, as well as the superoperator $\mathcal{D}$. For $\hat{H}_0$ and $\hat{\mu}_z$ the class **qm_operator** can be used. Since quantum mechanical operators can be expressed as Hermitian matrices, it is sufficient to store the real entries of the main diagonal as well as the complex entries of the upper triangle of the matrix. For example, a diagonal Hamiltonian

$$\hat{H}_0 = \begin{bmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{bmatrix} \tag{4.4}$$

can be created from a real vector $[\omega_1, \omega_2, \omega_3]$. In this case it is not necessary to provide a complex vector for the off-diagonal elements since they are set to zero by default. On the other hand, for a dipole moment operator

$$\hat{\mu}_z = \begin{bmatrix} 0 & \mu_{12} & \mu_{13} \\ \mu_{21} & 0 & \mu_{23} \\ \mu_{31} & \mu_{32} & 0 \end{bmatrix} \tag{4.5}$$

the main diagonal elements are defined by the real vector $[0, 0, 0]$ and the off-diagonal elements are represented by the complex vector $[\mu_{12}, \mu_{13}, \mu_{23}]$.

The treatment of superoperators is more complex, since the action of the superoperator on a quantum mechanical operator must be determined. Also, different choices of the superoperator are reasonable. Therefore, mbsolve features the class **qm_superoperator** and a derived sub-class **qm_lindblad_relaxation** that represents the resulting Lindblad superoperator. This approach allows future extensions, yet covers all current application examples, since the Lindblad form of the master equation is the most general Markovian description [34] (cf. Section 2.2.6). In Section 3.1 we have discussed the physical interpretation of the Lindblad dissipation superoperator and have identified the need for a compromise between usability and physical accuracy when pure dephasing terms are involved. In order to provide a physically accurate, yet practice oriented interface, mbsolve offers a user-friendly constructor for **qm_lindblad_relaxation**. It accepts a matrix

$$\begin{bmatrix} 0 & \gamma_{12} & \dots & \gamma_{1N} \\ \gamma_{21} & 0 & \dots & \gamma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N1} & \gamma_{N2} & \dots & 0 \end{bmatrix}, \tag{4.6}$$

where the γ_{ji} denote the relaxation (or scattering) rates from the basis states $|i\rangle$ to $|j\rangle$. As second parameter, the constructor of **qm_lindblad_relaxation** accepts a $N(N-1)/2$ real vector

$$[\gamma_{12,\mathrm{p}}, \gamma_{13,\mathrm{p}}, \gamma_{23,\mathrm{p}}, \gamma_{14,\mathrm{p}}, \dots, \gamma_{N-1,N,\mathrm{p}}] \tag{4.7}$$

with the pure dephasing rates, where the same ordering of off-diagonal elements as in **qm_operator** is used. The constructor tries to convert the dephasing rates into a corresponding coefficient matrix C (cf. Theorem 2.2.2) and checks whether this matrix is positive semi-definite. If the conversion or the check fails, a warning is emitted.

For the original Maxwell-Bloch equations, however, the general quantum mechanical description is unnecessarily complex. This form considers only two energy levels and a diagonal Hamiltonian $\hat{H}_0$, and usually neglects the static dipole moments ($\mu_{z,22} - \mu_{z,11} \approx 0$). For two energy levels, the master equation (3.69) reads

$$\begin{aligned} \partial_t \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} = &- \mathrm{i}\hbar^{-1} \left[\begin{bmatrix} H_{0,11} - \mu_{z,11}E & -\mu_{z,12}E \\ -\mu_{z,21}E & H_{0,22} - \mu_{z,22}E \end{bmatrix}, \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \right] \\ &+ \begin{bmatrix} -\gamma_{21}\rho_{11} + \gamma_{12}\rho_{22} & -\left[\frac{1}{2}(\gamma_{12} + \gamma_{21}) + \gamma_{12,\mathrm{p}}\right]\rho_{12} \\ -\left[\frac{1}{2}(\gamma_{12} + \gamma_{21}) + \gamma_{21,\mathrm{p}}\right]\rho_{21} & \gamma_{21}\rho_{11} - \gamma_{12}\rho_{22} \end{bmatrix}. \end{aligned} \tag{4.8}$$

Additionally, there is the constraint $\rho_{11} + \rho_{22} = 1$ on the populations, and for the coherence terms $\rho_{12} = \rho_{21}^*$ holds. Therefore, it is sufficient to determine the population inversion $w = \rho_{22} - \rho_{11}$, as the populations $\rho_{11} = (1 - w)/2$ and $\rho_{22} = (1 + w)/2$ can be derived from this quantity, and one of the coherence terms (usually ρ_{21}) [32]. Following this approach, and by assuming that the dipole moment $\mu_{z,21}$ is real, Eq. (4.8) can be brought into the form

$$\partial_t \rho_{21} = -\mathrm{i}\omega_{21}\rho_{21} - \mathrm{i}w\Omega_\mathrm{R} - \gamma_2\rho_{21}, \tag{4.9a}$$

$$\partial_t w = 4\Omega_R \Im\{\rho_{21}\} - \gamma_1 (w - w_0), \tag{4.9b}$$

which are the optical Bloch equations in the original form. Here, $\omega_{21} = \hbar^{-1}(H_{0,22} - H_{0,11})$ is the transition frequency between the two energy levels and $\Omega_\mathrm{R} = \hbar^{-1}\mu_{z,21}E_z$ is the instantaneous Rabi frequency. Furthermore, the dephasing rate

$$\gamma_2 = \frac{1}{2}(\gamma_{12} + \gamma_{21}) + \gamma_{12,\mathrm{p}}, \tag{4.10}$$

the scattering rate $\gamma_1 = \gamma_{12} + \gamma_{21}$, and the equilibrium population inversion

$$w_0 = \frac{\gamma_{21} - \gamma_{12}}{\gamma_{21} + \gamma_{12}} \tag{4.11}$$

are introduced to simplify the terms induced by the Lindblad superoperator.

In order to provide a convenient alternative for the Maxwell-Bloch equations in that form to the user, mbsolve features a subclass **qm_desc_2lvl** whose constructor accepts six real values. Those values represent the density $n_{3\mathrm{D}}$, the transition frequency, the dipole length $z_{21} = -e^{-1}\mu_{z,21}$ (where e is the elementary charge), the scattering rate, the dephasing rate, and the equilibrium inversion value, respectively. Then, the constructor builds the Hamiltonian

$$\hat{H}_0 = \frac{\hbar\omega_{21}}{2}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{4.12}$$

the dipole operator

$$\hat{\mu}_z = -ez_{21}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{4.13}$$

and the Lindblad superoperator. For the latter step, the scattering rate matrix

$$\begin{bmatrix} 0 & \gamma_{12} \\ \gamma_{21} & 0 \end{bmatrix} = \frac{\gamma_1}{2}\begin{bmatrix} 0 & 1 - w_0 \\ 1 + w_0 & 0 \end{bmatrix} \tag{4.14}$$

and the pure dephasing rate $\gamma_{12,\mathrm{p}} = \gamma_2 - \gamma_1/2$ have to be determined.

At this point, we have described the regions and materials of the device. Now we need to include the boundary conditions. As mentioned above, it is sufficient to store two real values that represent the reflectivity values of both ends of the device. Those values could be integrated directly into the device class. However, in order to maintain the flexible nature of our base library, the device class contains two pointers to an abstract class **bc_field**. Thereby, the project can be extended easily in future, e.g., to incorporate periodic boundary conditions. At the moment, the only subclass of the abstract class is **bc_field_reflectivity**, whose constructor accepts a real reflectivity value.

Scenario setup and initial conditions

As outlined above, the class **scenario** contains the dynamical part of the simulation setup. Namely, those are the source terms and the initial conditions. Although in most examples one source term is sufficient, the **scenario** can contain any number of terms, which are stored as pointers to the class **source**. Similar to other classes mentioned before, **source** is a base class that stores common information, such as the position x at which the source should be placed. Also, the **source** features a type field that distinguishes hard and soft sources. Here, it should be noted that a hard source sets the value of the electric field to the source value, whereas the soft source adds the source value to the current field value [117]. Different subclasses can be derived from the class **source**, such as **sech_pulse** and **gaussian_pulse**. As their names suggest, those subclasses yield a sech and a Gaussian pulse, respectively.

Similar to the treatment of the boundary conditions in the device, the scenario contains pointers to abstract classes that represent the initial conditions. Here, the pointer to **ic_density** specifies the initialization of the density matrix, and two pointers to **ic_field** determine the initial values of electric and magnetic field, respectively. Currently, only a subclass **ic_density_const**, which yields a constant initial density matrix, is implemented. For the fields there are more options: constant initialization, random initialization, and even a certain initial curve can be specified.

Apart from the simulation setup, further properties can be specified. For example, the number of spatial grid points can be specified. Thereby, the user can increase the accuracy and determine the effect on the results itself, as well as on the performance of the solver. Finally, the user needs to specify the desired results. Even in the most trivial simulations, several data sets are generated that are not required. In order to avoid wasting memory, a collection of **record** objects can be added to the scenario. Each record specifies a certain quantity that should be recorded, and contains information on the sampling interval and position. Then, during the simulation run, the solver analyses the information in the record list and stores the corresponding data traces in **result** objects. The latter are data container classes, which can be analyzed during postprocessing (either by accessing them in system memory or after exporting them to a file).

Solver and writer infrastructure

The base library only contains the abstract classes **solver** and **writer**, and leaves the implementation to libraries that build on the base. Before we discuss the resulting plugin structure in more detail, let us take a look at the common properties of all solvers and writers, respectively, that are represented by the abstract classes. The constructor of **solver** expects the name of the solver, as well as the device and the scenario to be simulated. After the solver is created, the method **run** executes the simulation. Then, the results can be extracted with the method **get_results**. The abstract class **writer** features a method **write** that accepts the results and writes them to a file. In addition to the results, the target filename, the device, and the scenario must be specified. The latter are required since the simulation result files should contain meta-information, such as the name of the device and the discretization size.

The plugin structure mentioned above guarantees the required flexibility. For the sake of brevity, this approach is described only for the solver. All remarks in the following hold analogously for the writer. The constructor of **solver**, which we already introduced, is indeed marked as protected. This means that instances of this class cannot be created directly. In order to create an instance of a certain solver, the static method **create_instance** must be called. This method expects the name of the solver as parameter (in addition to device and scenario), looks up the corresponding subclass of **solver**, and returns an instance of this subclass (using the provided device and scenario). While this approach may seem overly complex at first glance, it provides a clean interface to the user. For example, the user can acquire the available solvers with the static method **get_avail_solvers** and choose to create an instance of one of them without knowing the name of the corresponding subclass.

4.2.2 Writer for the Hierarchical Data Format (HDF5)

As already mentioned, the result objects can be passed to a writer. Currently, the only available writer implementation in mbsolve stores the data in the HDF5 format. This format is well accepted in computational science, and is supported by most programming languages (including C++, Python, MATLAB, and Octave) on all major platforms. The three main entities of the HDF5 format are groups, data sets, and attributes. Using groups, a hierarchical structure can be created. In each group (including the root), data sets and attributes can be placed, where attributes can be used to store meta-information.

The **writer-hdf5** stores the simulation meta-information (e.g., temporal discretization size) in attributes of the root group. Then, it creates a separate group for each result. Naturally, a data set containing the result data is added to this newly created group. However, since HDF5 does not natively support complex numbers, a second data set has to be added in case the result is complex (e.g., off-diagonal entries of the density matrix). Additionally, the writer creates a per-result attribute that informs the user whether the result is complex or real.

4.2.3 Project management and software quality assurance

From the very beginning of the mbsolve project, one goal of the project was to provide a reliable solution of high software quality to the scientific community. Indeed, any scientific software package must work reliably, as it serves as third pillar of science, the other two being theory and experiment. Clearly, creating a

reliable solution requires significant efforts, which is also one reason for making the resulting source code publicly available, in the hope that the efforts will help other research groups as well. As we shall see in Chapter 5, there are several best practices in scientific software engineering. The mbsolve project features the implementations of most best practices, which we discuss briefly in the following.

One elementary best practice is using a version control system. While it is beneficial even for a single user scenario, it becomes indispensable as soon as more developers work on the project. Ideally, it is combined with a collaboration web tool, which also offers an issue tracking system. The mbsolve source code is hosted on GitHub, which offers both. As to the quality of the source code, a coding convention has been defined and automated checks whether new contributions are compliant with it have been established. During the design of the software, the aim has been to provide a clear and object-oriented architecture. Modern features of the C++ programming language, such as smart pointers that prevent memory leaks, have been used in order to avoid typical mistakes, and to produce readable code. As far as third party components are concerned, open-source components have been selected exclusively. Additionally, the source code can be compiled with a variety of compilers (open-source and proprietary) on the three major operating systems Linux, Windows, and macOS. By using continuous integration (CI), the repetitive tasks (such as building the project, testing the resulting libraries, and performing additional tasks) have been automated. For example, the documentation is generated based on comments in the source code, compiled to static HTML pages, and uploaded to GitHub Pages [118].

4.2.4 Installation and requirements

As other researchers are invited to use and extend the mbsolve software, the build tools, the third party dependencies, and the steps required to build and/or install the software are described in the following.

In case the mbsolve software should be built from the source code, the CMake build system (version $\geq$ 3.9) and a recent C++ compiler are required. As to the latter, the mbsolve source code has been compiled successfully using the GNU Compiler Collection (gcc, version $\geq$ 4.9.2), the Intel C++ Compiler (any recent version), the Clang compiler (version $\geq$ v7.0.0), and the Microsoft Visual C++ compiler (MSVC, any recent version). It should be noted that AppleClang lacks OpenMP support entirely, and MSVC only supports a dated version, which may result in inferior performance. Then, the Eigen library (version $\geq$ 3.3.4) and the HDF5 libraries (any recent version) must be installed. The Python interface is optional, and requires Python (version $\geq$ 2.7) and SWIG (version $\geq$ 2.0.12). Finally, cxxopts (any recent version) is a prerequisite for the optional mbsolve-tool, and Doxygen (any recent version) is required for generating the documentation. If the required dependencies are not met, certain components may be disabled automatically by the build system.

Once the requirements are set up, the build process consists of executing CMake, which creates the project files for a certain generator, and running the generator, which depends on the platform in use (GNU make, Microsoft Visual Studio, etc.). The reader is referred to the documentation [118] for a more detailed description.

Alternatively, the compiled binaries can be installed directly. Here, the dependencies are a reduced set of the list above. Namely, those are the C/C++ standard libraries including OpenMP support, the Python runtime, and the HDF5 libraries. However, the same versions as used during compilation must be available, which is far from trivial. Therefore, it is recommended to install the binaries via conda, where the command

```
$ conda install -c conda-forge mbsolve
```

installs the mbsolve binaries together with all required dependencies. This approach works for all major platforms (Linux, Windows, and macOS), although the way to install a conda package may differ.

4.3 Application and simulation examples

In the following, the usage of the mbsolve project is demonstrated with the help of four application examples. Those examples have been selected so that they represent different simulation types (the Maxwell-Bloch

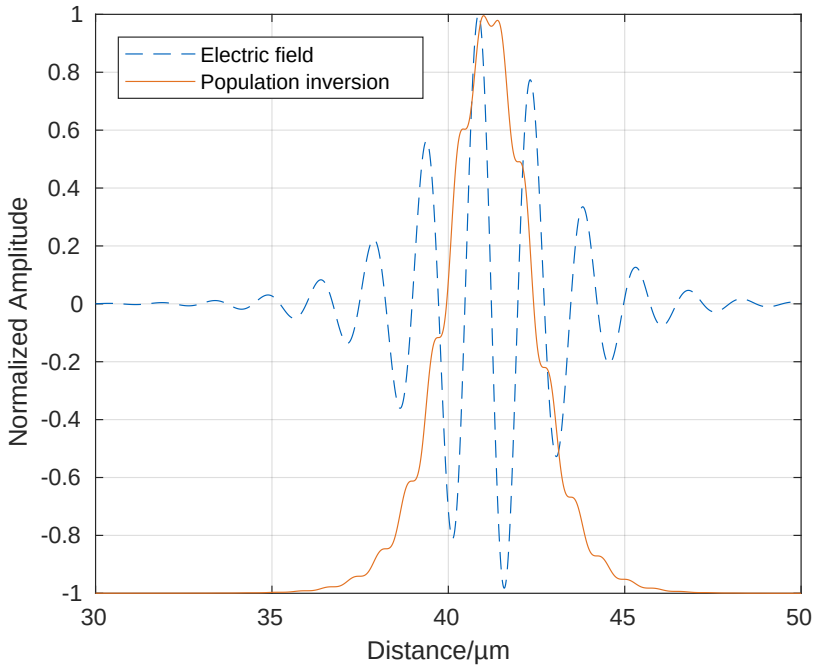


Figure 4.2 Simulation results from the self-induced transparency simulation setup, cf. Ziolkowski et al. [93], Fig. 2. Reprinted from M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations" [81] in accordance with the Elsevier Copyright Policy.

equations in 1D, or solving only the master equation at a certain point in space), use different features of mbsolve (different source types, initial conditions, etc.), and are executed using the C++ interface as well as the Python bindings. Furthermore, the selected simulations feature different numbers of energy levels, and while three examples are rather of theoretical nature, the final simulation models a real QCL. Apart from providing initial guidance to the prospective user, the examples serve as verification of the implementation.

4.3.1 Self-induced transparency (SIT) in two-level systems

The first application example reproduces the results presented in the work by Ziolkowski et al. [93]. In this pioneering work, the self-induced transparency (SIT) effect in two-level systems is investigated. Here, the active region is embedded in two short vacuum regions. In one vacuum region a sech pulse is injected and subsequently travels through the active region. By setting the pulse area to π, 2π, and 4π, the quantum mechanical systems are inverted once, twice, and four times, respectively. Figure 4.2 depicts the population inversion and the electric field for a 2π pulse, after it has propagated for 200 fs.

Listing 4.1 shows the complete Python code required to set up and run the SIT example. After the mbsolve libraries are imported, the script creates two materials, one of which features a simple two-level quantum mechanical description. The device to be simulated contains three regions, where vacuum is assigned to the first and third regions, and the active region is put in the middle. Initially, all quantum mechanical systems are inverted, which is represented by the initial density matrix `rho_init`. The scenario specifies 32768 spatial grid points, and that the simulation is run for 200 fs. Furthermore, two records are added that request the recording of the population inversion and the electric field, respectively. Both quantities are to be sampled using a time interval of 2.5 fs over the complete spatial domain. Also, a source, which represents a sech pulse at the left end of the device, is added to the scenario. Finally, the solver is created and runs the simulation. On a recent quad-core desktop computer this simulation should be processed in less than 20 seconds. The results are passed to the writer, which stores the data in a HDF5 file.

Listing 4.1 Python script for the self-induced transparency (SIT) setup, cf. `tools/python/ziolkowski1995.py` in the mbsolve repository [119].

```
# import mbsolve libraries
import mbsolve.lib as mb
import mbsolve.solvercpu
import mbsolve.writerhdf5

# vacuum
mat_vac = mb.material("Vacuum")
mb.material.add_to_library(mat_vac)

# simple two-level quantum mechanical description
# params: particle density, transition frequency,
#         transition dipole length,
#         scattering rate gamma_1, dephasing rate gamma_2,
#         equilibrium population inversion w_0
qm = mb.qm_desc_2lvl(1e24, 2 * math.pi * 2e14, 6.24e-11, 1.0e10,
                     1.0e10, -1.0)

# Ziolkowski active region material
mat_ar = mb.material("AR_Ziolkowski", qm)
mb.material.add_to_library(mat_ar)

# Ziolkowski device setup
dev = mb.device("Ziolkowski")
# params of each region: name, material, x_start, x_end
dev.add_region(mb.region("Vacuum␣left", mat_vac, 0, 7.5e-6))
dev.add_region(mb.region("Active␣region", mat_ar, 7.5e-6, 142.5e-6))
dev.add_region(mb.region("Vacuum␣right", mat_vac, 142.5e-6, 150e-6))

# initial density matrix
# params: lower level rho_11 fully populated, upper level rho_22
#         empty vector with coherence terms not set in params,
#         zero by default
rho_init = mb.qm_operator([ 1, 0 ])

# scenario
ic_d = mb.ic_density_const(rho_init)
ic_e = mb.ic_field_const(0.0)
sce = mb.scenario("Basic", 32768, 200e-15, ic_d, ic_e)
# record electric field and population inversion in 2.5fs intervals
sce.add_record(mb.record("inv12", 2.5e-15))
sce.add_record(mb.record("e", 2.5e-15))

# add source
sce.add_source(mb.sech_pulse("sech", 0.0, mb.source.hard_source,
                             4.2186e9, 2e14, 10, 2e14))

# run solver (advanced FDTD implementation + approach by
# Bidegaray 2001)
sol = mb.solver.create_instance("cpu-fdtd-red-2lvl-reg-cayley", dev,
                                sce)
sol.run()

# write results
wri = mb.writer.create_instance("hdf5")
outfile = dev.get_name() + "_" + sce.get_name() + "." +
          wri.get_extension()
```

```
results = sol.get_results()
wri.write(outfile, sol.get_results(), dev, sce)
```

4.3.2 Pulse propagation in a V-type three-level system

Using Maxwell-Bloch simulations, Song et al. [102] investigated a setup that is conceptually similar to the SIT example, the major difference being the active region. Here, atomic rubidium was modeled as a three-level quantum mechanical system. While the work mainly focused on the propagation of few-cycle pulses, it also contains the temporal simulation of a single three-level system that is driven by a pulse (cf. [102], Fig. 3). This case can be reproduced by solving the Lindblad master equation alone.

In Listing 4.2, the creation of a device and a scenario for this setup is given. Since a three-level system is considered here, it is necessary to use the complete quantum mechanical description including the operators and the superoperator. In this example, the device contains only one region of zero length. This is the first indication that the simulation does not consider any spatial dimensions. The second indication is the number of grid points, which is passed to the scenario constructor. Here, the number of spatial grid points is set to 1, whereas the 10000 temporal grid points are specified. Similar to the SIT setup, records and sources can be added, although it only makes sense to add them at the position of the single grid point. It should be noted that for a low number of spatial grid points parallelization of the calculation does not make sense. Fortunately, it is not needed here as this example will complete within seconds on a recent desktop computer (using only one core).

Listing 4.2 Code snippet of the Python script that reproduces the three-level driven quantum mechanical system in [102], cf. `tools/python/song2005.py` in the mbsolve repository [119].

```
# Hamiltonian
# (diagonal qm operator, vector for off-diagonal terms omitted)
energies = [ 0, 2.3717e15 * mb.HBAR, 2.4165e15 * mb.HBAR ]
H = mb.qm_operator(energies)

# dipole moment operator (qm operator with off-diagonal terms)
dipoles = [ -mb.E0 * 9.2374e-11, -mb.E0 * 9.2374e-11 * math.sqrt(2),
            0]
u = mb.qm_operator([ 0, 0, 0 ], dipoles)

# relaxation superoperator
rate = 1e10
# scattering rate matrix
rates = [ [ 0, rate, rate ], [ rate, 0, rate ], [ rate, rate, 0 ] ]
# pure dephasing rates are zero in this example
pure_deph = [ 0, 0, 0 ]
relax_sop = mb.qm_lindblad_relaxation(rates, pure_deph)

# initial density matrix
rho_init = mb.qm_operator([ 1, 0, 0 ])

# quantum mechanical description
qm = mb.qm_description(6e24, H, u, relax_sop)
mat_ar = mb.material("AR_Song", qm)
mb.material.add_to_library(mat_ar)

# Song setup
dev = mb.device("Song")
dev.add_region(mb.region("Active␣region␣(single␣point)", mat_ar, 0,
                         0))

# scenario
ic_d = mb.ic_density_const(rho_init)
```

```
ic_e = mb.ic_field_const(0.0)
ic_m = mb.ic_field_const(0.0)
sce = mb.scenario("Basic", 1, 80e-15, ic_d, ic_e, ic_m, 10000)
sce.add_record(mb.record("e", 0.0, 0.0))
sce.add_record(mb.record("d11", mb.record.density, 1, 1, 0.0, 0.0))
sce.add_record(mb.record("d22", mb.record.density, 2, 2, 0.0, 0.0))
sce.add_record(mb.record("d33", mb.record.density, 3, 3, 0.0, 0.0))

# add source
sce.add_source(mb.sech_pulse("sech", 0.0, mb.source.hard_source,
                             3.5471e9, 3.8118e14, 17.248, 1.76/5e-15,
                             -math.pi/2))
```

4.3.3 Six-level anharmonic ladder system

In the work by Marskar and Österberg [98], two simulation examples are discussed. The first example is a variation of the SIT setup in [93], which we have already discussed above. The second example considers the propagation of a Gaussian pulse in a medium that is modeled as a six-level anharmonic ladder system (cf. [98], Fig. 4). In this system, the energy levels are given as

$$\epsilon_{n+1} = \epsilon_n + \hbar\omega_0[1 - 0.1(n-3)], \tag{4.15}$$

where $n \in [1; N-1]$, N is the number of energy levels, and $\omega_0 = 2\pi \times 10^{13}\ \mathrm{s}^{-1}$ is the transition frequency between the energy levels ϵ_3 and ϵ_4. For convenience, and without loss of generality, $\epsilon_1 = 0$ is set to zero.

From our perspective, this setup is no more than a variation of the previous examples. However, it gives us a nice opportunity to introduce the mbsolve-tool, which features different simulation examples written in C++. In fact, all examples discussed in this section can be started by specifying the setup name as well as an appropriate solver and writer as command line arguments. For example, the anharmonic ladder simulation can be started as

```
$ mbsolve-tool -w hdf5 -m cpu-fdtd-red-6lvl-reg-cayley -d marskar2011-6lvl
```

Depending on the given setup name, the application creates the corresponding device and scenario, runs the specified solver, and uses the given writer to store the simulation results. Further command line arguments can be used to specify the number of spatial grid points and the simulation end time. As we shall see in Chapters 6 and 7, this feature is particularly handy for performance tests. By default, 8192 spatial grid points and a simulation end time of 2 ns are used, resulting in a runtime of approximately two minutes on a recent quad-core desktop computer.

We note that in the command line entry above, the number of energy levels seems to be stated explicitly as part of the device name. Indeed, the name "marskar2011-6lvl" and the shortcut "marskar2011" refer to the original setup, but in fact any number $N \geq 2$ can be specified. This feature has been added for performance comparisons of numerical methods, in which the performance is analyzed with respect to the number of energy levels (see Chapter 6). While the generalized example may not necessarily make sense from the modeling point of view, it represents a typical application example and hence constitutes a reasonable benchmark.

4.3.4 Quantum cascade laser frequency comb

Finally, we discuss the most complex and computationally demanding application example, as one simulation run required up to two hours on an AMD Ryzen Threadripper 2990WX machine using 16 cores. The quantum cascade laser frequency comb presented in [120] was modeled in previous work [23], where the experimental results were reproduced with good agreement. Most input parameters were determined using prerequisite Schrödinger-Poisson and ensemble Monte Carlo simulations. Thereby, the number of empirical model parameters was reduced to a minimum.

Listing 4.3 shows how a similar simulation can be set up using the mbsolve software. First, the quantum mechanical description of the active region material is created. Apart from the five eigenenergies on the main diagonal, the Hamiltonian in this example has non-zero off-diagonal elements, which account for tunneling effects. All the elements are determined by a Schrödinger-Poisson simulation in tight-binding basis [9]. The dipole moment operator is less spectacular as it only contains one non-zero element, which is placed on the off-diagonal element that corresponds to the transition between upper and lower laser level. The scattering rates and dephasing rates are used to set up the Lindblad relaxation superoperator. The quantum mechanical description and the active region material are then created using the parameters from [23]. Then, semi-transparent mirror boundary conditions are created, with the reflectivity values $R_1 = R_2 = 0.8$. The boundary conditions are subsequently used during the creation of the device, to which one region of the active region material is added. After that, the scenario is set up using the initial density matrix, in which only the upper laser level is populated. The electric field is initialized randomly to model spontaneous emission, whereas the magnetic field is set to zero. Since this is the default in mbsolve, we do not need to specify this choice of initial conditions explicitly. Finally, a record that triggers the recording of the electrical field at the facet is added to the scenario.

Listing 4.3 Code snippet of the mbsolve-tool C++ application that reproduces the quantum cascade laser frequency comb simulation in [23], cf. `mbsolve-tool/src/mbsolve-tool.cpp` in the mbsolve repository [119].

```
/* quantum mechanical description of active region */
/* params: vector of main diagonal entries, followed by vector of
 * off-diagonal entries */
mbsolve::qm_operator H(
  { 0.10103 * mbsolve::E0, 0.09677 * mbsolve::E0,
    0.09720 * mbsolve::E0, 0.08129 * mbsolve::E0,
    0.07633 * mbsolve::E0 },
  { 0.0, 1.2329e-3 * mbsolve::E0, -1.3447e-3 * mbsolve::E0,
    0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0 });

mbsolve::qm_operator u(
  { 0.0, 0.0, 0.0, 0.0, 0.0 },
  { 0.0, 0.0, 0.0, 0.0, 0.0, -mbsolve::E0 * 4e-9, 0.0, 0.0, 0.0,
    0.0 });

/* the scattering rate matrix */
std::vector<std::vector<mbsolve::real> > scattering_rates = {
  { 0.0000000, 0.4947e12, 0.0974e12, 0.8116e12, 1.0410e12 },
  { 0.8245e12, 0.0000000, 0.1358e12, 0.6621e12, 1.1240e12 },
  { 0.0229e12, 0.0469e12, 0.0000000, 0.0794e12, 0.0357e12 },
  { 0.0047e12, 0.0029e12, 0.1252e12, 0.0000000, 0.2810e12 },
  { 0.0049e12, 0.0049e12, 0.1101e12, 0.4949e12, 0.0000000 }
};

mbsolve::real deph_inj_ull = 1.0/(0.6e-12);
mbsolve::real deph_xxx_xxx = 1.0/(1.0e-12);

/* the vector of pure dephasing rates */
std::vector<mbsolve::real> dephasing_rates =
  { 0, deph_inj_ull, deph_inj_ull, deph_xxx_xxx, deph_xxx_xxx,
    deph_xxx_xxx, deph_xxx_xxx, deph_xxx_xxx, deph_xxx_xxx,
    deph_xxx_xxx };

auto relax_sop = std::make_shared<mbsolve::qm_lindblad_relaxation>(
    scattering_rates,
    dephasing_rates);

auto qm = std::make_shared<mbsolve::qm_description>(5.6e21, H, u,
                                                    relax_sop);
```

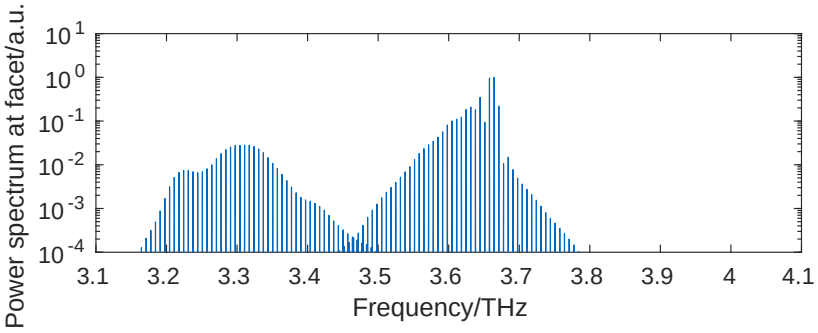


Figure 4.3 Power spectrum of the electric field recorded at the facet of the quantum cascade laser frequency comb. This simulation example bases on Tzenov et al. [23] (cf. Fig. 5c). For the corresponding experimental results cf. Burghoff et al. [120], Fig. 3b. Reprinted from M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations" [81] in accordance with the Elsevier Copyright Policy.

```
auto mat_ar = std::make_shared<mbsolve::material>("AR", qm, 12.96,
                                                  0.9, 1100);

/* set up device with semi-transparent mirror boundary cond. */
auto bc =
    std::make_shared<mbsolve::bc_field_reflectivity>(0.8, 0.8);
dev = std::make_shared<mbsolve::device>("tzenov2016", bc);
dev->add_region(std::make_shared<mbsolve::region>(
"Active region", mat_ar, 0, 5e-3));

/* initial density matrix (rho_33 is fully populated) */
mbsolve::qm_operator rho_init({ 0.0, 0.0, 1.0, 0.0, 0.0 });

/* basic scenario */
scen = std::make_shared<mbsolve::scenario>(
    "basic", num_gridpoints, sim_endtime, rho_init);
scen->add_record(std::make_shared<mbsolve::record>(
    "e0", mbsolve::record::electric, 1, 1, 0.0, 0.0));
```

After a solver has processed the device and scenario, the result data are written to a HDF5 file. At this point, we can discuss briefly the postprocessing of results. Typically, several tasks remain to be done after the solver has completed. In this case, the Fourier transform of the recorded electric field must be calculated and plotted. Naturally, this is beyond the scope of mbsolve and established software tools, such as MATLAB, Octave, or Python (with NumPy, SciPy, and Matplotlib), should be used. Examples for postprocessing MATLAB scripts can be found in `tools/matlab` in the mbsolve repository [119].

The simulation results (after postprocessing) are depicted in Fig. 4.3. While the spectrum of the electric field at the facet shows reasonable agreement with the experiment, there are features missing that the simulation in [23] could capture. This can be attributed to the differences in the underlying simulation methods. While the method in this work does not invoke the RWA, it does not account for chromatic dispersion or diffusion processes due to spatial hole burning.

4.4 Summary

In this chapter, we have reviewed the applications of the generalized Maxwell-Bloch equations in the one-dimensional form that is suitable for e.g., modeling various types of optoelectronic devices such as QCLs, where the wave guide geometry allows a reduction of the model to one spatial dimension. Besides

optoelectronic devices, the resulting equations can be applied to a variety of problems in which the plane-wave approximation is a reasonable assumption. Since numerical methods are required to solve these equations, we have assorted a list of requirements to a numerical solver tool and have found that there is no existing tool that fulfills all requirements.

Therefore, mbsolve, an open-source solver for the Maxwell-Bloch equations, has been presented in this chapter. The basic library of mbsolve provides a flexible and extensible framework to describe devices to be modeled, and simulation scenarios. On this basis, solvers can be written that implement different numerical methods and/or parallelization techniques. Similarly, writers can be implemented that export the simulation results to certain file formats, such as the open HDF5 file format. The resulting source code is written in C++ and features automatically generated bindings for Python. It is open-source and can be compiled using most of the established C++ compilers on all major platforms. Alternatively, the mbsolve software can be installed in binary form using the conda package manager.

The usage of mbsolve has been demonstrated with the help of four application examples, selected to cover different use cases. The first example features a two-level medium that can be described using a simplified version of the quantum mechanical description. The setup of the second simulation example can be considered a driven quantum mechanical system. In this simulation, there is no spatial coordinate, i.e., field propagation effects are not considered. The third example uses the generalized quantum mechanical description to handle the six-level medium. Finally, the last example is a simulation of an actual quantum cascade laser. This example is the most complex and computationally most demanding simulation.

Although the mbsolve software is already a helpful and reliable tool, several interesting issues are still unsolved, and potential optimization possibilities remain to be exploited. From the modeling point of view, the inclusion of dispersion in the electromagnetic properties should be one of the next steps. While it is not exactly trivial to implement, it is bound to play a significant role in the modeling of optoelectronic devices (see e.g., [121]). Also, the implementation of alternative boundary conditions will extend the application range of mbsolve, e.g., to the simulation of ring cavities. In order to enable numerically efficient simulations in two or three spatial dimensions, an attractive strategy could be to integrate the mbsolve code for the Lindblad equation into an established and high-performance open-source electromagnetics simulation project, such as MEEP.

The discussion of numerical methods is ongoing, as we shall see in Chapter 6. While we shall focus on methods that avoid the RWA/SVEA, we note that those approximations are beneficial for several applications. In future work, a solver that uses a numerical method that invokes the RWA (such as the Risken-Nummedal scheme [122]) could be implemented to support these applications better. The parallelization of the calculations shall be discussed in detail in Chapter 7. The mbsolve software already features solvers using the OpenMP standard for shared memory systems, but those existing solvers could be extended to distributed memory systems using Message Passing Interface (MPI) standard. Furthermore, offloading of calculations to graphics processing units (GPUs) is an intriguing possibility.

5 Ensuring quality in scientific software engineering

> Quality is to a product what character is to a man.
> — *Henry J. Heinz*

At this point, we take a short detour and review the measures required to obtain high-quality scientific software, based on the experiences made during development of the mbsolve software (see Chapter 4). As a recent essay in Nature [123] addressed, there is a familiar dilemma in (computational) science. On the one hand, science relies heavily on open-source software packages, such as libraries for mathematical operations, implementations of numerical methods, or data analysis tools. As a consequence, those software packages need to work reliably and should be easy to use. On the other hand, scientific software is notoriously underfunded and the required efforts are achieved as side projects or by the scientists working in their spare time.

Indeed, a lot of effort has to be invested beyond the work on the actual implementation – which is typically a formidable challenge on its own. This becomes apparent from literature on software engineering in general (such as the influential "Pragmatic Programmer" [124]), and in scientific contexts in particular (e.g., [125]–[128]). The vast number of best practices guides and development guidelines available (e.g., those published by the German Aerospace Center (DLR) [129] and the Netherlands eScience Center [130]) further underlines the importance of the topic and may serve as guidance, but often scientists lack the time and/or formal training in software engineering required to ensure sustainable software development [123], [127], [128]. Too often, this results in poorly maintained software projects of questionable reliability and usability.

Given all this, once again the goal is to achieve much with little effort. Therefore, in this chapter a project skeleton is presented that may serve as solid yet lightweight base for a small to medium-scale scientific software project. In the envisaged use case, scientists can create an instance of this template in just a few clicks. This instance implements essential best practices in software engineering from the very start. After performing a minimal number of customizations, the scientist can soon start working on the actual implementation and can concentrate on what really matters. In the context of pure C++ [131] or Python [132] projects, such skeletons have already proven their value. The focus of this chapter is on scientific software libraries which are written in the C++ programming language for performance reasons and feature bindings for Python in order to provide an easy-to-use interface to the user, such as the mbsolve software. It should be noted that the skeleton concentrates on this particular use case and does not (and should not) cover every eventuality (e.g., when support for the Fortran programming language is required). This is contrary to the recommendations in related literature, which are kept general and language-agnostic on purpose. The rationale behind this decision is to keep the template lightweight and avoid cluttering.

This chapter is organized as follows: In Section 5.1, we identify the essential best practices that are required to ensure high-quality scientific software based on related literature and the experiences with the mbsolve software project. Subsequently, the project skeleton and the specific implementation of the measures identified are presented in Section 5.2. As already stated above, some minor customization steps are required. Section 5.3 gives an overview of these steps and thereby an introduction to the (potential) user. Finally, we conclude with a short summary and give an outlook on future work, i.e., additional tools and measures that further improve the quality of scientific software projects.

This chapter is largely based on M. Riesch, T. D. Nguyen, and C. Jirauschek, "bertha: Project skeleton for scientific software", *PLOS ONE*, vol. 15, no. 3, e0230557, Mar. 2020. DOI: 10.1371/journal.pone.0230557, an open access article published under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

5.1 Best practices in scientific software engineering

This section describes the essential recommendations and best practices from related literature [123]–[130] that serve as basis for the project skeleton. All recommendations are language-agnostic and grouped into seven categories with no particular order of importance. Table 5.1 gives an overview of the best practices.

5.1.1 Project management

Most software projects in a scientific context start with a single developer. However, over time the projects are likely to grow, be extended, and possibly taken over by other developers. Building a developer community is crucial for the success of the project in general and in particular for open-source projects [126]. Therefore, the project infrastructure should be able to handle multiple developers from the very start.

All of the guidelines in literature mention the usage of a version control system (VCS). This is beneficial even for the single developer, as a VCS intrinsically features a backup solution and synchronization between different machines. Once more developers start working on the project, the VCS enables transparent collaboration. By using a VCS, the "Make Incremental Changes" paradigm [127], [128] can be implemented easily and the intrinsically generated development history may serve as rudimentary documentation of design decisions [126].

In a more advanced scenario, the VCS is coupled with a project management tool that provides a means of communication within the developer team, and thereby further enhances transparency. As the communication logs are available for developers who join the team at a later stage, this also provides a certain form of documentation [126]. One essential element of a project management tool is a ticket system or issue tracker. Issues are requests for a certain change (such as a bug fix or feature implementation) and play a crucial role in modern iterative and incremental software development processes, such as feature-based development [133]. As the name suggests, issue trackers keep track of issues from their creation (by users or developers) to their completion in the form of an accepted solution by the developer [129]. Modern project management tools also include convenient mechanisms for code review. Similar to a scientific paper, a rigorous review process may be time-intensive and annoying, but eventually yields solutions of higher quality and wider acceptance [127].

5.1.2 Code quality

Just as we care about language style when writing a scientific article, so we should care about coding style when writing scientific software. Here, we should bear the mottos "Write Programs for People, Not Computers" [127] and "Don't Repeat Yourself" [124], [127] in mind and produce easily readable and modular code. In developer teams, it is crucial to agree on a certain coding style at the beginning of the project. The coding style usually consists of two parts: rules for formatting code and best practices for programming in the respective language. Code formatting tools enable manual and automated checks to establish whether the source code is compliant with agreed code formatting rules [130]. Analogously, static code analysis tools check whether the agreed best practices are violated [129].

5.1.3 Independence

Some guidelines recommend that open standards, protocols, and file formats should be used wherever possible (e.g., the HDF5 format for large data sets [130]). Thereby, vendor lock-in situations are avoided which would arise, for example, if a certain source code can only be compiled using a certain compiler brand or version. The general recommendation here is to provide solutions that work with the most widely used operating systems and compilers (and possibly combinations thereof) from the very start.

Following the advice that one should never reinvent the wheel, established software libraries and tools are often used to speed up development processes. Here, we recommend using open-source components unless there is a strong reason not to. This is in agreement with the interoperability and reusability part of the FAIR principle [134], [135].

Table 5.1 Overview of best practices in software engineering for scientific software projects. For each best practice, implementation candidates are listed where the selected choice is denoted in bold.

Group	Best practice	Implementation candidates (not exhaustive)
Project management	Version control system	**git**, mercurial, svn
	Project management tool	**GitLab**, **GitHub**, Bitbucket, JIRA
	Workflow	**GitLab Flow**, GitHub Flow, git flow
Coding style	Code formatting style	**Mozilla**, LLVM, Google, Chromium
	Code formatting tool	**clang-format**
	Static code analysis	clang-tidy, cppcheck, cpplint
Independence	Use open file formats	e.g., JSON, CSV, HDF5
	Use open-source libraries	e.g., Eigen, FFTW, GNU Scientific Library
Automation	Continuous integration	**gitlab-ci**, **Travis CI**, AppVeyor, Microsoft Azure
	Build automation	**CMake**, GNU make, Bazel, Ninja, MS Build
Documentation	Function reference	**Doxygen**, Sphinx (with Breathe)
	"Big picture" documentation	**Markdown**, reStructuredText
Testing	Unit test framework	**Catch2**, Google Test, Boost Test Library
	Code coverage report	**gcov**, various commercial tools
Deployment	Package binaries	**conda**, Conan, Debian apt
	Online documentation	**GitLab Pages**, **GitHub Pages**, readthedocs.io

5.1.4 Automation

We should "Let the Computer Do the Work" [124], [127] and automate repetitive tasks such as building the software, running tests, performing quality checks, and deploying the generated artifacts (typically, software in binary form and documentation) to a software repository. Otherwise, those tedious tasks are most likely postponed, not done at all, or performed only partially. Here, continuous integration (CI) tools are helpful as different jobs can be defined and grouped into stages, which are executed every time the developers push changes to the version control repository. Then, the developers receive feedback on the changes, which is an essential part of the "Make Incremental Changes" strategy [127].

The feedback typically consists of (at least) two parts, which are briefly outlined. First, the build process should run in an automated and platform-independent fashion. Here, it is particularly important that third-party dependencies are found without hard coded paths. The output of the build process tells the developers whether the build on different platforms was successful. This is especially beneficial as most developers develop on a certain platform and the code is not intrinsically tested on other platforms (different operating systems, different compiler versions, etc.). Second, test programs can be executed automatically on different platforms. For example, unit tests can help to verify the correct behavior of certain functions or modules of the software. Functional tests, on the other hand, help to gain more confidence in the overall function of the software [129].

It makes sense to define the continuous integration pipelines as early as possible, so that the developers benefit from the feedback from the very beginning. Thereby, bugs in the software (in particular regressions) can be detected early. Furthermore, the effectiveness of optimizations can be assessed while the correct operation of the software is ensured.

5.1.5 Documentation

In order to make scientific software reusable, providing documentation to users and developers is one of the most important steps [123]–[130]. Bangerth and Heister [126] list five items that the documentation should contain: traditional comments, function level documentation, class level documentation, overview of how modules interact, and complete examples in tutorial form. As to traditional comments, it is good practice to "Document Design and Purpose, Not Mechanics" [127] and avoid obvious comments. Function and class level documentation is typically generated based on comments in code using special annotation. The resulting reference manual is particularly interesting for developers and advanced users who need to know the details. On the other hand, the module overview documentation should inform new users about the big picture. This information is typically written into the files README (aim of the software, installation notes, list of dependencies), CHANGELOG (overview of releases, features, known bugs), CODE_OF_CONDUCT and CONTRIBUTING (guidelines for (potential) developers), as well as TUTORIAL (guide for (potential) users) [128].

5.1.6 Testing

As mistakes are natural and are bound to happen, we should plan for them and develop strategies on how to detect them as early as possible [127]. Automated testing, the importance of which has already been underlined in Section 5.1.4, is the cornerstone of such strategies. It should be noted that the effectiveness of tests should be monitored as well. Here, code coverage tools are useful as they are able to detect code parts which are not covered by the executed tests [129].

Again, we stress that certain measures, such as writing unit tests, should be carried out from the very beginning. Apart from their use in automated testing, unit tests may have a positive effect on the code design. Since modular code is usually testable, performing unit tests can be considered a necessary requirement for modular code [124].

5.1.7 Deployment

Whether or not a certain software project is used depends to a large degree on the ability to distribute it [126]. Hence, it is advisable to package the software and distribute it using an established software repository [123]. Similar to the practices discussed above, it is important that the deployment is carried out automatically and as early as possible [125].

5.2 Implementation of the project skeleton

Based on the (general and language-agnostic) best practices introduced in the section above, measures for a C++ software library with bindings for the Python language are implemented in this section. The result is publicly available [136] and may serve as a template for new projects or reference for existing projects. Figure 5.1 outlines the skeleton approach.

It should be noted that there may be different ways to implement a certain measure. For the sake of simplicity, we discuss only one or two possibilities for most measures. Following the recommendations in Section 5.1.3, open-source tools and libraries have been selected exclusively. Thereby, one particular lightweight solution is provided for scientists who are new to the topic, while the advanced users may replace a certain implementation of a measure with another library or tool of their choice.

Since a project skeleton does not include a real implementation, best practices regarding planning, structuring, and writing code can hardly be demonstrated. In this regard, the reader is referred to available literature on the topic, such as [124].

5.2.1 Usage of a version control system (VCS) and appropriate workflow

A multitude of version control systems has been published and used over the last three decades. We stick to our criterion that the software must be open-source and note that git has received much attention since it was first released in 2005. It features distributed version control and a flexible branching model, rendering it perfectly suited for open-source projects. However, the flexible branching model might, at the same time, be a significant drawback. Each project should define a workflow to show how changes are developed, tested, and integrated. As usual, it makes sense to use something established, such as the GitLab Flow [138]. This workflow uses feature branches to develop and test new features or bug fixes. Once the changes on the feature branch fulfill the requirements and pass the automated tests and quality checks, the developer can open a merge request. A maintainer can subsequently merge the changes in the main development branch. Additionally, the GitLab Flow allows stable branches and different environments (such as production) in which further restrictions may apply. The latter features are not required at the initial stage of a project, but underline that the GitLab Flow is simple enough for small projects yet powerful enough for large and established projects.

5.2.2 Usage of a project management tool including issue tracking

There are several management tools and hosting platforms that can be combined with the git version control system with different strengths and drawbacks. Here, we can leave the choice to the developers and provide two possible solutions for the undecided.

Over the last decade, the GitHub platform has received significant attention. It provides free public git repositories and integrations with other services (such as the zenodo repository for storing research output). Due to its prominence, a mirror repository of the project skeleton is available in GitHub [139]. This repository is marked as project template, which allows a new project to be instantiated with a few clicks. As to continuous integration, GitHub offers support for external CI providers such as Travis CI, AppVeyor, and Microsoft Azure. These services are typically free for open-source projects and configured using a YAML file, where CI jobs can be described. As an example, a basic configuration file for Travis CI has been added to the skeleton. It triggers build and unit tests on Linux, Windows, and macOS platforms given that the user registers on Travis CI, where the CI operation can be activated for the repository in question.

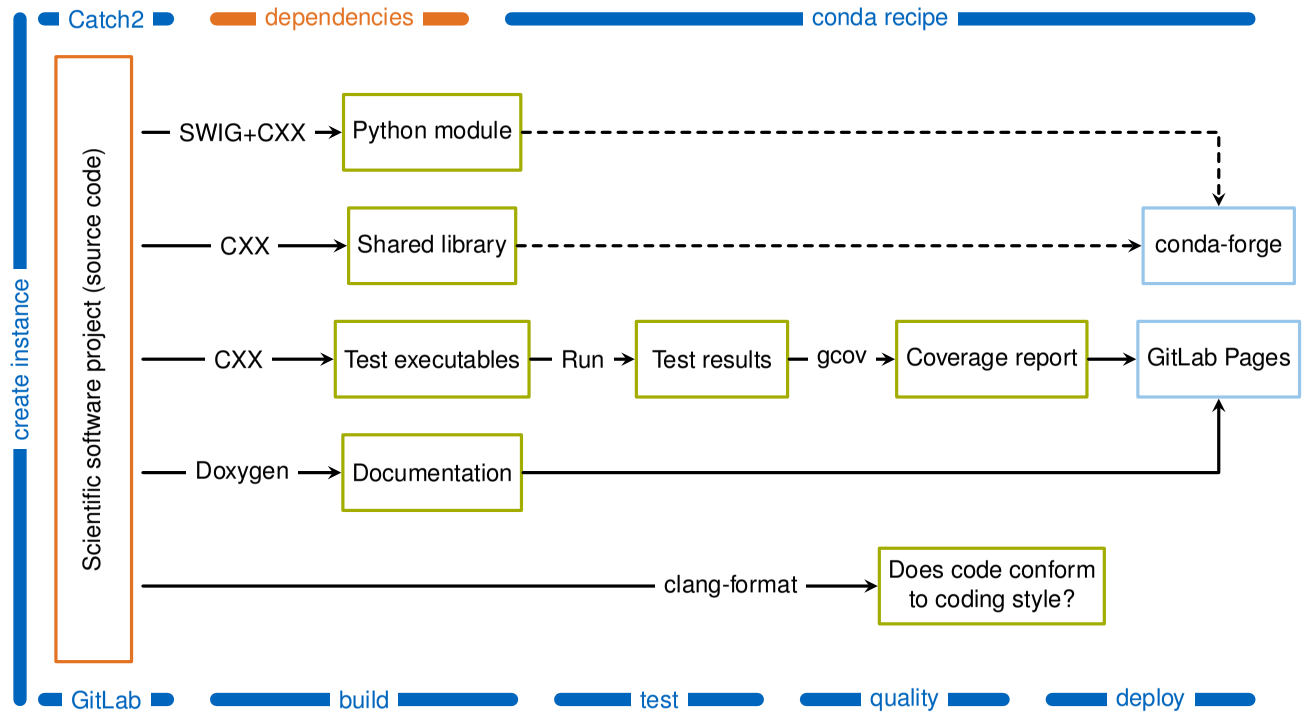


Figure 5.1 Overview of the project skeleton. The source code and dependencies of a scientific software project are denoted in orange. These are the parts the developer has to provide. The presented skeleton guides the project from creation to deployment. Here, the arrows denote jobs that are created by the CMake build system. These jobs are triggered during the different continuous integration stages (build, tests, quality, deploy) or (in the case of the dashed arrows) by the conda-forge build service that follows the recipe [137]. The job names indicate the tools in use, where CXX represents one of the C++ compilers that are supported by CMake. Reprinted from M. Riesch *et al.*, "bertha: Project skeleton for scientific software" [116] (CC BY 4.0).

Alternatively, the GitLab platform can be used, which is conceptually similar to GitHub, the main difference being the possibility of self-hosting the platform on a local server for free. While the concepts (such as Pages and Releases) are similar, there are slight differences. For example, the project template instantiation mechanism is different. At this point, it is not possible to create an instance of the project skeleton with a single click. However, this feature could be provided in the near future [140].

GitLab.com provides free hosting and internal continuous integration services for open-source projects. Currently, those internal CI services are restricted to the Linux operating system. It is possible, however, to install GitLab's CI suite on a local machine and connect it to GitLab. Alternatively, an external service can be used for Windows or macOS operating systems. In the event that the project should not be open-source, the self-hosted operation mode may be selected. Here, the CI suite must be installed on local machines, which can subsequently be connected to the local GitLab installation.

It should be noted that the skeleton does not contain configuration files for all options in order to provide a lightweight solution. Instead, the configuration file for the GitLab internal CI has been added, which calls the targets generated by the build systems. From this configuration file, corresponding files for other CI services can be derived.

5.2.3 Automated build system

In particular when the C++ programming language is involved, the CMake project provides well-established tools to build, test, and package software. The main advantage of CMake (compared to alternatives such as GNU make, Visual Studio, or Eclipse) is that a level of abstraction is introduced. The configuration files consist of directives such as `add_library` or `find_package` and are, therefore, quite easy to read and understand. Based on those configuration files, project files for the aforementioned alternatives (and many other build systems) can be generated. Thereby, the software project can be built for different operating systems or using different compilers. In addition, CMake features a mechanism for finding third-party libraries and tools. This feature is essential for cross-platform dependency management.

As a proof of concept, a simple shared library written in C++ has been added to the project skeleton. It features a simple class `device` with two member variables that represent its start and end coordinates, respectively. An instance of this class can be created using one of two constructors, where either the coordinates are specified directly, or the length can be set and the start coordinate is assumed to be at the origin. Finally, a method returns the length of the device.

For such a shared library, Python bindings can be generated conveniently using the SWIG project. It is fully supported by CMake and requires only a minimal configuration file, which basically specifies which C++ header files should be considered when the interface is created. SWIG scans the specified header files and automatically generates a Python module, which can be subsequently imported and used in a Python project.

5.2.4 Unit testing

Ideally, the software is designed so that each unit of software (e.g., a function) fulfills a certain, unique task ("Design by Contract" technique [124]). Furthermore, the implementation of each unit is flawless. While the first goal can be achieved by careful design and refactoring, the second statement is rarely true. As mentioned above, mistakes will happen and we have to test whether the implementations of each unit work correctly.

In the case of our simple C++ library, we have to check, for instance, whether the calculation of the length yields the correct result. This can be achieved by writing a unit test that creates an instance of the `device` class, calls its `get_length` method, and compares the result of the method to the expected value. Also, whenever the user specifies input data, the implementation should check whether those values are reasonable and deal with invalid values (most likely, by throwing an exception). Error handling code must be tested as well, for example by creating a unit test in which the error is provoked on purpose and checking whether the error handling code yields the correct behavior. As the number of unit tests is expected to be large for a real life project, it is recommended to use a unit test framework.

Here, the Catch2 library is chosen as it is open-source, lightweight and header-only. Based on this library, a test executable with several unit tests has been added to our CMake build system using the CTest functionality. Whether or not the unit tests cover all possible situations can be assessed using code coverage tools. The possibility of using the gcov tool has been added to the project skeleton. This tool generates profiling information during the execution of tests. This information can be subsequently converted to a human-readable report, in which metrics such as line coverage are given on a per-file basis.

5.2.5 Automatic code formatting

Here, the clang-format tool constitutes a helpful and versatile instrument. It can be configured using a single file, in which the code formatting rules are specified. There are several predefined styles that can be used as-is, or alternatively serve as a basis. It is also possible to define a certain style from scratch, but it is recommended to use an existent style (with slight modifications, if required).

In the project skeleton, the clang-format tool is integrated into the CMake build system, making it easy for the user to format all source files automatically. This functionality is also used to check whether the source code conforms to the specified style in the scope of continuous integration.

5.2.6 Documentation generation

From the implementation point of view, we can separate the different types of documentation listed in Section 5.1.5 into two groups, namely the function reference and the overview documentation. The function reference is based on comments in the source code that use special annotation. The information in those comments can be extracted using the Doxygen tool. For the overview documentation, which provides the "big picture", it makes sense to use a structured text format. Since Doxygen supports the Markdown language, files such as README.md and CONTRIBUTING.md are written in this annotation. Both the overview documentation and function reference are then transformed into static HTML pages that can be viewed locally or uploaded to a web server.

We note that while Doxygen provides unchallenged support for in-source C++ documentation, the design of the generated HTML files appears a bit dated. More advanced workflows are available that use Doxygen as input parser and alternative tools to generate the static HTML pages. However, this is beyond the scope of the work at hand.

5.2.7 Automated packaging and deployment to a public repository

While many operating systems or programming languages feature a common repository for exchanging programs and libraries in binary form, it would be beneficial to have a language-agnostic repository that covers all operating systems. Fortunately, the conda system provides exactly this. Once a software project is in a stable state, a recipe can be created on conda-forge that defines the source of the project, the steps required to build it, and meta information such as the name of the responsible maintainer. Based on this recipe, the conda-forge build system automatically generates the binaries for different platforms. Then, on each platform the resulting package can be easily installed within a conda environment.

Most likely, the package has dependencies on other libraries. The conda system offers a vast number of third-party components and convenient methods of installing them. The environment approach already mentioned has a positive effect on the dependency management, as in Windows it is generally impossible to distinguish between different versions of a library (at least when considering unmanaged C++ code), dubbed the "DLL Hell". Using conda environments, however, it is possible to separate different versions in a clean and convenient way.

The documentation generated could be included in a conda package as well. However, it is usually more appropriate to publish it on a web server for visibility reasons. Both GitHub and GitLab offer the possibility of hosting static HTML pages, such as those generated by Doxygen. With a few lines of CI configuration, the documentation is automatically generated and uploaded. See [141] for an example.

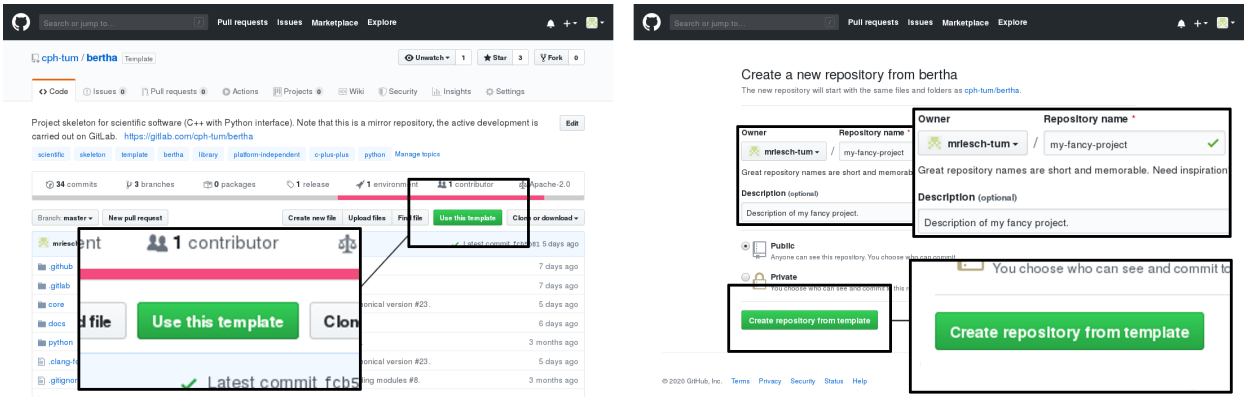


Figure 5.2 Creating an instance of the project skeleton on GitHub. On the project page of bertha [139], click on "Use this template". In the following, enter the desired owner, repository name, and project description. The button "Create repository from template" will then create the instance. Reprinted from M. Riesch *et al.*, "bertha: Project skeleton for scientific software" [116] (CC BY 4.0).

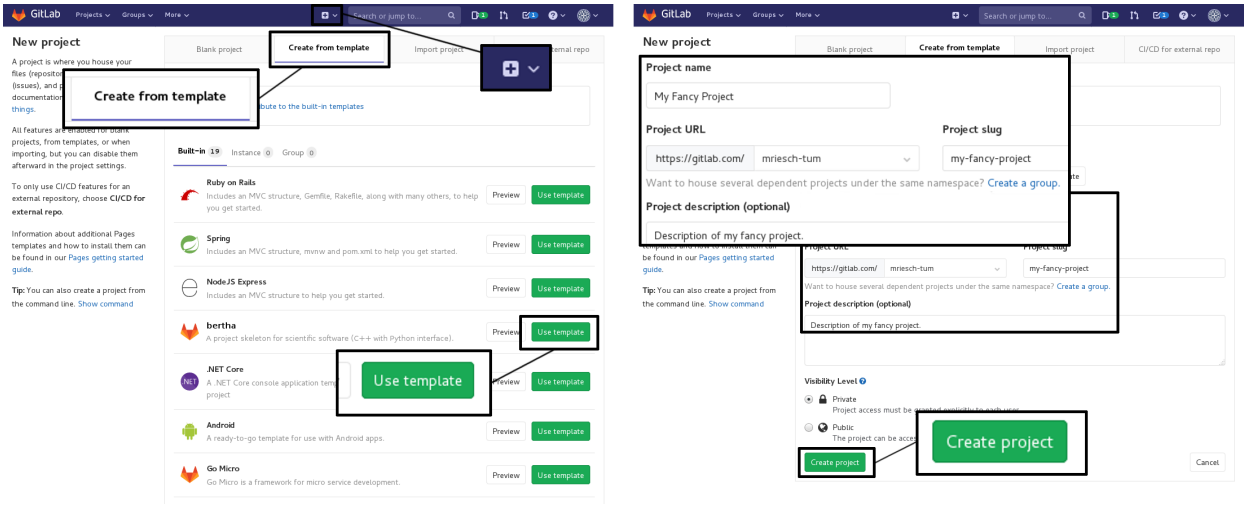


Figure 5.3 Creating an instance of the project skeleton in GitLab. Click on the plus button to create a new project. After selecting the "Create from template" tab, choose bertha by clicking "Use this template" (currently in development, see [140]). Then, enter the project name and description and click "Create project". Reprinted from M. Riesch *et al.*, "bertha: Project skeleton for scientific software" [116] (CC BY 4.0).

5.3 Creating a skeleton instance

In order to create a new project, the project skeleton can be cloned using the mechanisms of either GitHub (see Fig. 5.2) or GitLab (as described in Fig. 5.3). Alternatively, the files can be copied manually and added to a new repository. After the cloning procedure, the skeleton can be adjusted to the needs of the new project. The steps recommended and required are outlined briefly in the following. For more detailed instructions, the reader is referred to the "Tutorial" section in the bertha documentation [141]. It should be noted that registration on Travis CI and activation of the project is required for continuous integration support in GitHub.

5.3.1 Setup stage

At the beginning, it is important to define a meaningful name for the project and replace bertha with this name throughout the project (e.g., in the CMake build structure). It should be ensured that the name is not already used (e.g., in conda-forge) if the project is to be open-source. Then, the project team should agree on where to host the project (for internal use only or publicly available), on the license for the project, and

on the workflow. The latter includes mainly the coding style and the version control workflow. Both should be documented as soon as possible.

5.3.2 Implementation stage

At this point, the software project has a solid initial state. Now it is time to add functionality. Here, the developers should consider writing the documentation first (the contract), then implementing the functionality, and at the same time writing unit tests. This approach will seem slow but improves the quality of the design and helps to detect mistakes early on. Also, the CMake build structure can be adjusted to add requirements (e.g., software libraries) or additional modules (besides the existing core library).

5.3.3 Publication stage

In the case of an open-source project, the code should be distributed and communicated as soon as it has some first functionality. For the distribution of the project in binary form, the conda recipe for bertha [137] may serve as reference.

5.4 Summary

In this chapter, a skeleton for scientific software projects has been presented, which consist of libraries written in the C++ programming language and feature a Python interface. The skeleton contains the essential elements required to ensure best software engineering practices. This project aims to provide the scientific community with a helpful tool that saves time during the setup of a new project. Based the experience gained during the development of the skeleton, creating a bertha instance may replace at least one person month of evaluating tools, reading documentation, and searching for answers in the internet.

Furthermore, this contribution may serve as checklist and reference for existing projects. It is envisaged that in both use cases – building a project from scratch and adapting an existing one – the project skeleton shall aid the implementation of good practices in scientific software engineering and consequently improve the quality and reusability of scientific software projects.

As a next step, the implementation of further measures is envisaged. For example, a static code analysis tool could further improve the quality of the code. Also, the generated documentation and quality reports should be presented with a modern appearance. Finally, the project skeleton concept could be transformed to other project classes in scientific software engineering, such as a combination of a Fortran library with a Python interface. In this regard, there is already one adaptation of the skeleton for scientific document projects written in LaTeX [142], [143].

6 Numerical treatment of the Maxwell-Bloch equations

Bender's Computer Dating Service – Discreet and Discrete
— *Futurama*

Smile, you buggers! Pretend it's Christmas!
— *Pete Townshend, The Who at the Isle of Weight*

As already mentioned in the previous chapters, the generalized Maxwell-Bloch equations can usually only be solved numerically. The required numerical treatment is discussed in this chapter, whose objective is to find accurate and efficient numerical methods that are suitable for the long-term simulation of QCL devices in frequency comb or ultra-short pulse operation. We shall see in Section 8.1 that invoking the rotating wave approximation (RWA) in an ultra-short pulse generation setup requires some a priori knowledge of the central frequency. Full-wave simulations, on the other hand, solve the same problem in self-consistent fashion. Also, since the RWA assumes that the electromagnetic radiation features a narrow spectrum and small peak intensities it will break down eventually in the simulation scenarios mentioned above. Therefore, we focus on numerical methods beyond the RWA in the following. For a review on numerical methods that invoke the RWA the interested reader is referred to [32]. The attribute long-term stems from the fact that simulations of frequency combs often require a relatively large simulation end time in order to acquire the radiation spectrum in high spectral resolution. Here, it is crucial that the numerical methods preserve the physical properties in general and the properties of the density matrix in particular. As reported in [105], [144], some methods may produce unrealistic results such as negative populations.

With these requirements in mind, we review the numerical methods for Maxwell's equations in the scope of the generalized Maxwell-Bloch equations in Section 6.1. Additionally, we discuss the coupling between Maxwell's equations and the quantum mechanical master equation and thereby provide the basis to integrate the numerical methods for the master equation, which are in the focus of Section 6.2. Then, Section 6.3 describes the implementation of the most promising candidates and analyzes their single-thread performance. After a summary of the findings, a brief outlook on possible future developments concludes this chapter.

6.1 Numerical methods for Maxwell's equations

Out of the many methods that solve Maxwell's equations numerically, mainly two are used in the context of Maxwell-Bloch equations. The majority of related work (e.g., [93], [94], [96], [97], [105]) uses on the finite-difference time-domain (FDTD) [117] method to solve Maxwell's equations. Clearly, it is one of the standard approaches in computational electromagnetics, and its simplicity allows straightforward implementation of source terms and sharp material boundaries. As disadvantages, the need for a constant

Section 6.1 is largely based on Section 2.2 of M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations", *Comput. Phys. Commun.*, vol. 268, p. 108 097, Nov. 2021. DOI: 10.1016/j.cpc.2021.108097 in accordance with the Elsevier Copyright Policy.

Section 6.2 is largely based on M. Riesch and C. Jirauschek, "Analyzing the positivity preservation of numerical methods for the Liouville-von Neumann equation", *J. Comput. Phys.*, vol. 390, pp. 290–296, Feb. 2019. DOI: 10.1016/j.jcp.2019.04.006 in accordance with the Elsevier Copyright Policy.

spatial discretization and its numerical dispersion must be mentioned. The need of the FDTD method for a constant spatial discretization can lead to inefficient discretization patterns in setups with different media. The grid spacing must be chosen to suit the medium with the largest effective refractive index and will be unnecessarily small for media with smaller refractive indices. The usual remedy for the numerical dispersion is to decrease the spatial and temporal grid spacing, which in turn increases the computational workload. In practice, spatial discretization sizes between $\lambda/20$ and $\lambda/200$ have been used, where λ denotes the smallest occuring wavelength [32].

As an alternative, the pseudo-spectral time-domain (PSTD) method has been used in related work [98], [107]. The PSTD method calculates the spatial derivatives accurately in a pseudo-spectral domain, in which they are reduced to mere multiplication operations. Thereby, the spatial contribution of the numerical dispersion is eliminated and the temporal contribution is significantly reduced. This relaxes the requirement for the spatial discretization size, since (in theory) only the Nyquist-Shannon theorem needs to be considered. In practice, the discretization size $\lambda/10$ has been found to achieve reasonable accuracy [98]. However, this comes at the cost of potentially expensive calls to the fast Fourier transform (FFT), and complex implementations of boundary conditions and source terms.

We note that numerical methods that adapt the spatial discretization variably to the problem are used to solve Maxwell's equations, but that they have not been applied in the scope of the Maxwell-Bloch equations. Also, although the PSTD method provides an intriguing alternative, the efforts to implement and extend it are estimated to outweigh its benefits. Therefore, we follow the majority of related literature and focus on the FDTD method in the remainder of this thesis.

6.1.1 The finite-difference time-domain method

We begin the numerical treatment of Maxwell's equations (4.1) and (4.2) with the discretization of the variables therein. To this end, we introduce the spatial index m and discretization size Δx as well as the temporal index n and discretization size Δt. Then, the continuous variables, such as the electric field $E_z(x,t)$, are sampled at discrete grid points in space and time, so that $E_z^{m,n} = E_z(m\Delta x, n\Delta t)$. At this point, the Yee grid [145] comes into play as the central element of the FDTD method. The main feature of the Yee grid (depicted in Figure 6.1) is that the electric and magnetic field are staggered in time and space. This clever choice allows the usage of central difference schemes for the spatial and temporal derivatives, which feature second-order accuracy [117]. Using this discretization, the differential equations (4.1) and (4.2) can be transformed into the difference equations

$$\begin{aligned}\frac{E_z^{m,n+1} - E_z^{m,n}}{\Delta t} =& -\epsilon^{-1}\sigma\frac{E_z^{m,n+1} + E_z^{m,n}}{2} - \epsilon^{-1}\Gamma\left(\partial_t P_z\right)^{m,n+1/2} \\ &+ \epsilon^{-1}\frac{H_y^{m+1/2,n+1/2} - H_y^{m-1/2,n+1/2}}{\Delta x}\end{aligned} \tag{6.1}$$

and

$$\frac{H_y^{m+1/2,n+1/2} - H_y^{m+1/2,n-1/2}}{\Delta t} = \mu^{-1}\frac{E_z^{m+1,n} - E_z^{m,n}}{\Delta x}. \tag{6.2}$$

While the treatment of the derivatives is straightforward [147], the terms $-\sigma E_z$ and $-\Gamma\partial_t P_z$ deserve further mention. In general, care is taken that the right-hand side of the differential equations are evaluated at a half step on the temporal discretization, thus allowing a leapfrog integration scheme [117]. Since for the term $-\sigma E_z$ this is not directly feasible, the value $E_z^{m,n+1/2}$ is approximated by averaging between the old value $E_z^{m,n}$ and the new value $E_z^{m,n+1}$. As a consequence, Eq. (6.1) becomes an implicit equation. However, we can readily rewrite it as an explicit update equation. The polarization term $(\partial_t P_z)^{m,n+1/2}$ is evaluated at the desired position by definition, but needs to be discussed in detail in the next section.

With regard to stability of the FDTD method, the Courant-Friedrichs-Lewy condition

$$\Delta t < \frac{n}{c}\Delta x \tag{6.3}$$

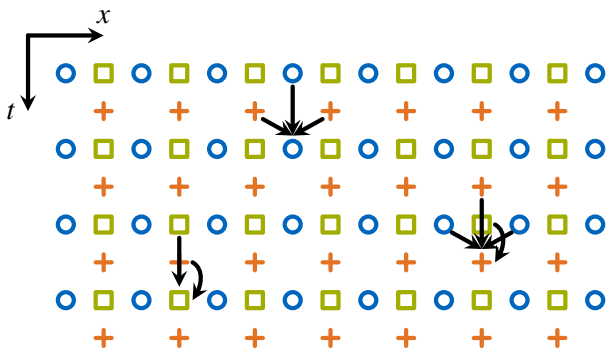


Figure 6.1 The standard Yee grid of the FDTD method combined with the discretization of the density matrix with respect to time and space. Electric and magnetic field are denoted with orange crosses and blue circles, respectively. The density matrix discretization is marked using green squares. The arrows indicate the data dependencies during the update of three quantities. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

must be fulfilled, where $c/n = (\mu\varepsilon)^{-1/2}$ is the speed of light in the material under consideration. Since the speed of light may vary in a simulation setup with different materials, the largest value for c/n must be selected, resulting in a minimal Δt. Equation (6.3) can be rewritten as

$$\Delta t = C\frac{n}{c}\Delta x, C < 1, \tag{6.4}$$

where the Courant number C has been introduced [117].

6.1.2 Coupling of the master equation to Maxwell's equations

With the discretization of the electromagnetic field in mind we can advance to the discussion of the remaining components of the Maxwell-Bloch equations, namely the polarization calculation in Eq. (3.68) and the evolution of the density matrix, which is governed by Eq. (3.69). We begin with the calculation of the polarization term

$$\begin{aligned}\partial_t P_z &= n_{3\mathrm{D}} \operatorname{Tr}\left\{\hat{\mu}_z \partial_t \hat{\rho}\right\} = n_{3\mathrm{D}} \operatorname{Tr}\left\{-\mathrm{i}\hbar^{-1}\hat{\mu}_z\left[\hat{H}_0 - \hat{\mu}_z E_z, \hat{\rho}\right] + \hat{\mu}_z \mathcal{D}(\hat{\rho})\right\} \\ &= n_{3\mathrm{D}} \operatorname{Tr}\left\{-\mathrm{i}\hbar^{-1}\hat{\mu}_z\left[\hat{H}_0, \hat{\rho}\right] + \hat{\mu}_z \mathcal{D}(\hat{\rho})\right\},\end{aligned} \tag{6.5}$$

where we use the master equation (3.69), the property

$$\left[\hat{H}_0 - \hat{\mu}_z E_z, \hat{\rho}\right] = \left[\hat{H}_0, \hat{\rho}\right] - \left[\hat{\mu}_z E_z, \hat{\rho}\right] \tag{6.6}$$

of the commutator, and the properties (cyclic property and linearity) of the trace operation, which lead to the relation

$$\operatorname{Tr}\left\{\hat{\mu}_z\left[\hat{\mu}_z E_z, \hat{\rho}\right]\right\} = E_z \operatorname{Tr}\left\{\hat{\mu}_z\hat{\mu}_z\hat{\rho}\right\} - E_z \operatorname{Tr}\left\{\hat{\mu}_z\hat{\rho}\hat{\mu}_z\right\} = 0. \tag{6.7}$$

The resulting equation can be discretized in straightforward fashion and then reads

$$\left(\partial_t P_z\right)^{m,n+1/2} = n_{3\mathrm{D}} \operatorname{Tr}\left\{-\mathrm{i}\hbar^{-1}\hat{\mu}_z\left[\hat{H}_0, \hat{\rho}^{m,n+1/2}\right] + \hat{\mu}_z \mathcal{D}\left(\hat{\rho}^{m,n+1/2}\right)\right\}. \tag{6.8}$$

In this choice of discretization (depicted in Fig. 6.1), the density matrix is sampled at the same temporal discretization points as the magnetic field, which is referred to as weak coupling. A different approach is to sample the density matrix at the same temporal discretization points as the electric field (strong coupling). While both approaches are stable in principle, methods that use weak coupling tend to be easier and computationally more efficient, as the density matrix and the electric field can be updated alternatingly using explicit update equations [106].

Regardless of the coupling, the remaining task is the update step of the density matrix from one time value to the next one. This operation can be formally written as

$$\hat{\rho}^{m,n+1/2} = \mathcal{V}\left[\Delta t\,(n+1/2), \Delta t\,(n-1/2), \hat{\rho}^{m,n-1/2}\right] \approx \mathcal{V}^{m,n}\left(\hat{\rho}^{m,n-1/2}\right), \tag{6.9}$$

where $\mathcal{V}$ is the time evolution superoperator of the reduced density matrix (cf. Section 2.2.6). In practice, however, this formal expression can only be approximated using a certain update superoperator $\mathcal{V}^{m,n}$, which serves as placeholder for different numerical methods. We note that the (otherwise time-independent) update superoperator is likely to depend on the electric field E_z, which raises the question at which spatial discretization points the field is sampled. Since the electric field value $E_z^{m,n}$ is readily available in the scope of the weak coupling approach, we aim to use this value and discretize the update superoperator at the very same discretization points.

6.2 Completely positive trace-preserving methods for the master equation

In the previous section, we have introduced the update superoperator $\mathcal{V}^{m,n}$ as placeholder for the numerical method that solves the master equation (3.69). As next step, we review the state of the art in numerical methods and discuss suitable candidates to replace this abstract placeholder. Several methods have been presented in related literature and can be divided into different categories. Ziolkowski, Slavcheva and coworkers used a Crank-Nicolson (CN) scheme in their work, where the implicit nature of the CN scheme was resolved with a predictor-corrector (PC) approach [93], [96]. In the following, we will treat these two methods separately. The methods used in [98], [106], [107], [148], [149] solve Eq. (3.69) exactly by calculating matrix exponentials. This category shall be referred to as ME methods. Finally, several implementations of the Runge-Kutta (RK) method have been presented (see e.g. [94], [97], [108]).

As pointed out above, the numerical methods must preserve the properties of the density matrix. We recall from Section 2.2.6 that this is guaranteed if the update superoperator $\mathcal{V}^{m,n}$ is a completely positive trace-preserving (CPTP) map. Bidégaray et al. [105] analyzed the positivity preservation of the CN scheme and found that the update step of this method does not necessarily yield a positive semi-definite density matrix. In a subsequent publication, Songolo and Bidégaray stated that the Runge-Kutta method does not preserve the properties of the density matrix but no rigorous proof was given [150]. Finally, while the ME methods preserve the properties by definition in theory, the actual implementations must be analyzed since several approximation technique are applied in practice. Therefore, we analyze in the following whether or not the exact ME approaches, the ME approaches using approximations, the Runge-Kutta method, and the predictor-corrector technique feature a CPTP update step [144]).

As we shall see, it may be beneficial to switch to a vector representation of the density matrix. The columns of the $N \times N$ density matrix $\hat{\rho}$ can be stacked on top of each other into a vector $\boldsymbol{\rho}$ with N^2 complex elements. This column-major order is denoted as $\boldsymbol{\rho} = \mathrm{col}(\hat{\rho})$ and called representation in Liouville space [98]. The master equation (3.69) then reads $\partial_t \boldsymbol{\rho} = (L + D)\,\boldsymbol{\rho}$, where L and D are $N^2 \times N^2$ matrices representing the Liouvillian $\mathcal{L}$ and the dissipation superoperator $\mathcal{D}$, respectively.

6.2.1 Matrix exponential approaches that solve the equation exactly for a single time step

The numerical methods of this group calculate the solution $\mathcal{V}^{m,n} = \exp[(\mathcal{L}^{m,n} + \mathcal{D})\Delta t]$ in order to solve the differential equation exactly at every time step n. The form of this exponential (and therefore the form of the update map) is determined by the representation. For example, in Liouville space this update reads

$$\boldsymbol{\rho}^{m,n+1} = \exp\left[(L^{m,n} + D)\,\Delta t\right]\boldsymbol{\rho}^{m,n} =: V^{m,n}\boldsymbol{\rho}^{m,n}, \tag{6.10}$$

where $V^{m,n}$ is the matrix that represents the update superoperator $\mathcal{V}^{m,n}$ in Liouville space. It should be noted that the Liouvillian is assumed to be time-independent during an update step, which is perfectly reasonable in the scope of Maxwell-Bloch simulations with the FDTD method.

This update map is completely positive and trace preserving by definition, since the differential equation is solved analytically at every time step and the form of the master equation guarantees a CPTP map as solution

(cf. Section 2.2.6). Here, we assume that there are numerical methods that solve the matrix exponential with machine precision so that the CPTP conditions will not be altered by the implementation. Obviously, such methods exist and are implemented in publicly available libraries such as the Eigen library [151]. Here, the scaling and squaring method combined with the Padé approximation [152] is implemented which calculates the exponential of a $M \times M$ matrix in $O(M^3)$ time. Since in Liouville space the matrices are $N^2 \times N^2$, this approach has the complexity $O(N^6)$, which is of course a significant drawback. Therefore, the algorithm presented in [153] is a promising alternative. It exploits the fact that not the matrix exponential itself but its action on a vector is asked. This action can be determined in $O(N^4)$ time and up to a user-defined accuracy. Other algorithms base on Krylov subspace methods [154], [155] such as the approaches presented in [156], [157]. A similar technique is described in [149], [158], [159], where the matrix exponential function is expressed using Chebychev polynomials. It should be noted that while the algorithm using the action of the matrix exponential is designed for the application in Liouville space only, the other methods may be used in both Liouville space and regular representation. Then, the performance of the methods may differ depending on the representation. The reason for this is the dimension of the Liouville space, which represents a large hindrance that may only be overcome if the involved matrices are sparse.

6.2.2 Matrix exponential approaches that use approximations

The evaluation of the matrix exponential function is costly, in particular when the matrices in Liouville space are concerned. Therefore, related work focused on the solution of the Liouville-von Neumann equation in regular representation. A closed analytic expression such as in Eq. (6.10) cannot be derived, but as remedy the symmetric Strang operator splitting technique [160] can be invoked [98], [105]–[107]. This approach splits the exponential

$$\mathcal{V}^{m,n} = \exp\left[(\mathcal{L}^{m,n} + \mathcal{D})\,\Delta t\right] \approx \exp\left[(\mathcal{L}_1 + \mathcal{D})\,\Delta t/2\right] \exp(\mathcal{L}_2^{m,n}\Delta t) \exp\left[(\mathcal{L}_1 + \mathcal{D})\,\Delta t/2\right] \tag{6.11}$$

into two parts that constitute the solution to the time-independent part $\mathcal{L}_1 + \mathcal{D}$ of the master equation's right-hand side, and to the time-dependent Liouvillian $\mathcal{L}_2^{m,n} = \mathrm{i}\hbar^{-1}[\hat{\mu}_z E_z^{m,n}, \hat{\rho}]$, respectively. Both solutions can usually be determined separately. This splitting produces an error of order $O(\Delta t^2)$ (except in the unlikely case where both parts of the Liouvillian commute). However, if each part of the Liouvillian yields a CPTP map as solution, the combination is again a CPTP map.

The separation of the time-dependent part has the advantage – apart from allowing analytic solutions – that the time-independent solution has to be determined only once and can be precalculated. Hence, related literature focused on the efficient evaluation of the time-dependent solution

$$\exp\left(\mathcal{L}_2^{m,n}\Delta t\right)\hat{\rho} = \exp\left(\mathrm{i}\hbar^{-1}\mu_z E_z^{m,n}\Delta t\right)\hat{\rho}\exp\left(-\mathrm{i}\hbar^{-1}\mu_z E_z^{m,n}\Delta t\right), \tag{6.12}$$

as this evaluation must be performed at every time step.

Since the exponential of an $N \times N$ matrix $\mu_z E_z^{m,n}$ can be achieved with a complexity of $O(N^3)$ [152] and the multiplication of $N \times N$ matrices completes in $O(N^{\approx 2.37})$ time [161], the exponential calculation dominates the complexity of the update step. The goal of the approaches outlined below is to approximate the exponential using matrix multiplications and sums. Then, the complexity is reduced and sparse algorithms may be used (note that the exponential of a matrix is usually dense, even when the matrix itself is sparse).

In [105], [106] the exponential $\exp(\mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t) =: \hat{A}$ is approximated using the Crank-Nicolson scheme. We can readily see that this exponential is the solution of the simple differential equation $\partial_t \hat{A} = \mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\hat{A}$. The CN scheme is applied to this differential equation (in contrast to the work in [93], where the CN scheme is applied to the master equation) and the approximation

$$\exp\left(\mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t\right) \approx \left(\hat{I} - \mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t/2\right)^{-1}\left(\hat{I} + \mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t/2\right) \tag{6.13}$$

is derived. This approximation can be related to the Cayley transform of the skew-Hermitian matrix $\mathrm{i}\hat{\mu}_z$, which is guaranteed to be unitary [162]. Thus, we can readily see that the conditions of Theorem 2.2.1 are still fulfilled. Of course, an additional numerical error is introduced, but the density matrix properties

are preserved thanks to the clever choice of the approximation. In terms of complexity, the matrix inverse operation and the matrix multiplication are equal [163] and the complexity of the update step is $O(N^{\approx 2.37})$. In [150], a variation of this approach was described in the context of nonstandard finite difference methods, where the resulting schemes are currently limited to the elementary but essential case with two energy levels. In this case, analytic solutions exist for the general matrix exponential and further simplifications may be applied, e.g., by assuming that the main diagonal entries of $\hat{\mu}_z$ are zero.

Other approaches (e.g., [164]) use the Taylor series to evaluate the matrix exponential. According to [152], this method converges slowly and will therefore show inferior performance or relatively large numerical errors. As an alternative, one could think of an approximation

$$\exp(\mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t) \approx \hat{I} + \mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t - (\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t)^2/2 + \cdots + (\mathrm{i}\hbar^{-1}\hat{\mu}_z E_z^{m,n}\Delta t)^k/k! =: \hat{B} \quad (6.14)$$

based on the truncated Taylor series. This way the update step has the form $\hat{B}\hat{\rho}_n\hat{B}^\dagger$ as required by Theorem 2.2.1, but the condition in Eq. (2.81) is not fulfilled since $\hat{B}\hat{B}^\dagger = \hat{I} + O(\Delta t^l)$, where $l = 2\lfloor 1 + k/2 \rfloor$. Hence, this technique does not feature a CPTP update step.

6.2.3 Runge-Kutta method

Several research groups [94], [97], [108] used the fourth-order Runge-Kutta method (see e.g., [147]) to solve the master equation. Here, the update step reads

$$\mathcal{V}^{m,n}\hat{\rho} = \hat{\rho} + \Delta t\,(k_1 + 2k_2 + 2k_3 + k_4)/6, \quad (6.15)$$

where $k_1 = \mathcal{F}^{m,n}(\hat{\rho})$, $k_2 = \mathcal{F}^{m,n+1/2}(\hat{\rho} + \Delta t k_1/2)$, $k_3 = \mathcal{F}^{m,n+1/2}(\hat{\rho} + \Delta t k_2/2)$, and $k_4 = \mathcal{F}^{m,n+1}(\hat{\rho} + \Delta t k_3)$, and $\mathcal{F}^{m,n} = \mathcal{L}^{m,n} + \mathcal{D}$ denotes the right hand side of the master equation. This method is promising since the computational workload of the update step is dominated by multiplications of $N \times N$ matrices (assuming that we apply the Runge-Kutta method in regular representation). Similar to the matrix exponential methods using approximations above, the computational complexity is $O(N^{\approx 2.37})$ and sparse methods can be applied.

We apply this method to a simple test system with the Liouvillian $\mathcal{L}(\hat{\rho}) = -\mathrm{i}\hbar^{-1}[\hat{H}, \hat{\rho}]$, where the Hamiltonian $\hat{H}$ is time-independent, and transform the update step to the Liouville space. In this representation, the vector $\boldsymbol{\rho}$ is updated using

$$\boldsymbol{\rho}^{m,n+1/2} = \left[I + L\Delta t + \frac{1}{2}(L\Delta t)^2 + \frac{1}{6}(L\Delta t)^3 + \frac{1}{24}(L\Delta t)^4\right]\boldsymbol{\rho}^{m,n-1/2} =: V^{m,n}\boldsymbol{\rho}^{m,n-1/2}, \quad (6.16)$$

where I is the $N^2 \times N^2$ identity matrix and the Liouvillian $L = \mathrm{i}\hbar^{-1}(\hat{H}^* \otimes \hat{I} - \hat{I} \otimes \hat{H})$. Here, $\hat{I}$ is the $N \times N$ identity matrix, the tensor product $\otimes$ is the Kronecker product in this case, and the asterisk denotes the complex conjugate. See [98] for a detailed description of the Liouvillian in Liouville space.

Now we rewrite the update matrix $V^{m,n}$ using $\tilde{H} = -\mathrm{i}\hbar^{-1}\Delta t\hat{H}$ and assess whether it can be decomposed into

$$V^{m,n} = \sum_{j=0}^{\infty} c_j \sum_{k=0}^{j} \frac{j!}{k!\,(j-k)!}\left(\tilde{H}^*\right)^{j-k} \otimes \tilde{H}^k \stackrel{!}{=} \sum_{i=1}^{N^2} \hat{V}_i^* \otimes \hat{V}_i, \quad (6.17)$$

which is the condition in Eq. (2.80) in Liouville space [165]. For the Runge-Kutta method, the coefficients c_j on the left-hand side are zero for $j > 4$. Also, it becomes apparent that if such matrices $\hat{V}_i$ exist, they must be functions of the Hamiltonian $\tilde{H}$. We assume that Δt is chosen sufficiently small so that we can expand each matrix $\hat{V}_i = \sum_{l=0}^{\infty} a_{i,l}\tilde{H}^l$ as Taylor series. Then, the decomposition reads

$$\sum_{i=1}^{N^2} \hat{V}_i^* \otimes \hat{V}_i = \sum_{i=1}^{N^2}\sum_{l=0}^{\infty}\sum_{k=0}^{\infty} a_{i,l}^*\left(\tilde{H}^*\right)^l \otimes a_{i,k}\tilde{H}^k = \sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\sum_{i=1}^{N^2} a_{i,l}^* a_{i,k}\left(\tilde{H}^*\right)^l \otimes \tilde{H}^k. \quad (6.18)$$

We note that since $c_6 = 0$ for the Runge-Kutta method, for $j = 6$ and $k = 3$ the corresponding term in Eq. (6.17) vanishes, and consequently the term with the powers $l = k = 3$ in Eq. (6.18) should also vanish.

Therefore, the sum $\sum_{i=1}^{N^2} |a_{i,3}|^2 = 0$ and subsequently all coefficients $a_{i,3}$ must be zero. However, a term with the powers $j = 3$ and $k = 0$ is present in the update matrix in Eq. (6.17), but the corresponding term (with $l = 3$ and $k = 0$) in Eq. (6.18) vanishes if all $a_{i,3}$ are zero. We deduce that the update matrix cannot be decomposed and the update map of the Runge-Kutta method is not CPTP.

6.2.4 Predictor-corrector technique

In their work, Ziolkowski, Slavcheva et al. [93], [96] treat the master equation with the Crank-Nicolson scheme. The positivity preservation of this scheme has been discussed in [105]. However, the actual implementation uses the predictor-corrector technique to resolve the implicit nature of Crank-Nicolson. Hence, we shall concentrate on the explicit predictor-corrector method in the following.

The predictor-corrector update step begins by setting $\hat{\rho}_{\mathrm{PC}} = \hat{\rho}^{m,n-1/2}$ and then executes the procedure

$$\hat{\rho}_{\mathrm{PC}} \leftarrow \hat{\rho}^{m,n-1/2} + \frac{\Delta t}{2} \mathcal{F}^{m,n} \left(\hat{\rho}_{\mathrm{PC}} + \hat{\rho}^{m,n-1/2} \right) \tag{6.19}$$

four times, where $\mathcal{F}$ again denotes the right-hand side of the master equation. Then, the result is assigned to the value $\hat{\rho}^{m,n+1/2} = \hat{\rho}_{\mathrm{PC}}$. Again, we consider a simple test system with a time-independent Liouvillian and write for the complete update step in Liouville space

$$\boldsymbol{\rho}^{m,n+1/2} = \left[I + L\Delta t + \frac{1}{2} (L\Delta t)^2 + \frac{1}{4} (L\Delta t)^3 + \frac{1}{8} (L\Delta t)^4 \right] \boldsymbol{\rho}^{m,n-1/2} =: V^{m,n} \boldsymbol{\rho}^{m,n-1/2}, \tag{6.20}$$

We can readily see that apart from different coefficients c_j the predictor-corrector technique and the Runge-Kutta method have the same update step. Therefore, we can deduce that the computational complexity is the same (the predictor-corrector method can be implemented in regular representation as well and sparse methods can be applied) and the update map of the predictor-corrector technique is not CPTP. Indeed, by using the argumentation above, one can show that no method of the form

$$\boldsymbol{\rho}^{m,n+1/2} = \left[\sum_{j=0}^{M} c_j (L\Delta t)^j \right] \boldsymbol{\rho}^{m,n-1/2} \tag{6.21}$$

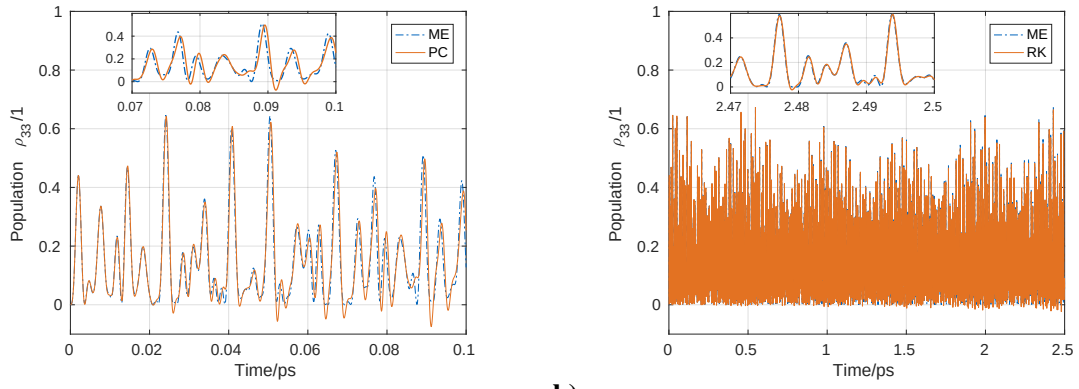
with a finite number of steps M can be decomposed to fulfill the condition in Eq. (2.80).

6.2.5 Verification

As we have shown above, only one group of the typically used numerical methods – namely, the matrix exponential methods – can be represented as completely positive trace preserving map and is therefore guaranteed to yield realistic results in long-term simulations. However, it remains to be demonstrated that the results of this theoretical analysis are relevant in a practical example. Hence, we implement a simple simulation based on related literature and compare the results of the different numerical methods. We consider again the anharmonic ladder example from Section 4.3.3 [98], since it is a well-established, multi-level setup that constitutes a significant challenge for the numerical methods in terms of accuracy and performance. Similar to our simple test system above, we only consider a time-independent Hamiltonian (i.e., the interaction term $-\hat{\mu}_z E_z$ and the dissipation superoperator $\mathcal{D}$ are ignored).

The Runge-Kutta method and the predictor-corrector technique are implemented in MATLAB, as well as a matrix exponential approach that serves as reference. By using the variable-precision arithmetic (vpa) toolbox of MATLAB, the precision can be increased and it can be assured that the results below are not affected by round-off error artifacts. All methods used the same time step size $\Delta t = 0.1$ fs, which was chosen sufficiently small in order to avoid stability issues. The resulting MATLAB scripts are publicly available as open-source project [166].

Figure 6.2 depicts the simulation results of all three methods. The matrix exponential method solves the problem exactly and serves as reference. By close inspection we can see that the population ρ_{33} remains in



a) b)

Figure 6.2 Simulation results of the anharmonic ladder system using **a)** the predictor-corrector approach and **b)** the Runge-Kutta method. The results of the matrix exponential method serve as reference in both parts of the figure. For details of the curves see the respective inset. Reprinted from M. Riesch and C. Jirauschek, "Analyzing the positivity preservation of numerical methods for the Liouville-von Neumann equation" [144] in accordance with the Elsevier Copyright Policy.

the interval $[0, 1]$, as the theory dictates. In contrast, the results of the predictor-corrector method (Fig. 6.2a) show that after a short duration the population becomes negative, which is clearly a violation of the properties of the density matrix. Similarly, the population becomes negative when using the Runge-Kutta method (see Fig. 6.2b). In this case, the first instance of a negative population occurs at a later point in time, which is consistent with the order of accuracy of the methods (a fourth-order Runge-Kutta method was used, the predictor-corrector approach is a second-order method). Nevertheless, the Runge-Kutta method may yield unrealistic results for certain simulation end times. Finally, it should be noted that although the population ρ_{33} was chosen as figure of merit, the discussed features are also visible in other populations.

6.3 Implementation and single-thread performance comparison

Following the review of numerical methods for Maxwell's equations and the master equation in the sections before, we discuss in this section the implementation of the most promising candidate in the scope of the mbsolve solver tool. This method is then used as a reference to benchmark two alternative methods, whose implementation details are also in the focus of this section. It has been pointed out frequently that solving the master equation is the major bottleneck in Maxwell-Bloch simulations [98], [106], [107], [109]–[111], [167]. Therefore, the aim of this section is to evaluate different numerical approaches for the master equation, which are weakly coupled to the same FDTD implementation for Maxwell's equations.

All three approaches discussed in the following subsections belong to the group of matrix exponential methods, which guarantee a CPTP update map. The section is concluded by a performance comparison of these three approaches.

6.3.1 Operator splitting in regular representation (os-reg-cayley)

The method by Bidégaray et al. [105], [106] is selected as the established reference method. It uses the operator splitting technique described in Eq. (6.11) to divide the right-hand side of the master equation into a time-independent part, which contains the contributions of the Hamiltonian $\hat{H}_0$ and the dissipation superoperator $\mathcal{D}$, and a time-dependent part, which can be related to the interaction term $[\hat{\mu}_z E_z, \hat{\rho}]$. Then, assuming a diagonal Hamiltonian $\hat{H}_0$ solutions for both parts can be readily derived. In the scope of this work, however, we aim to avoid any assumptions about the Hamiltonian $\hat{H}_0$ and use a slightly different splitting approach. To this end, we divide the right-hand side of the master equation into the Liouvillian $\mathcal{L}$, which contains the Hamiltonian $\hat{H} = \hat{H}_0 - \hat{\mu}_z E_z$, and the dissipation superoperator. In [106] a symmetric

splitting is suggested, in which the solution $\exp(\mathcal{L}\Delta t)$ of the non-dissipative, Liouvillian part for a full time step is embedded by two solution operators $\exp(\mathcal{D}\Delta t/2)$ that solve the dissipative part for a half time step each. The complete update operator then reads

$$\mathcal{V}^{m,n} = \exp\left(\mathcal{D}\Delta t/2\right) \exp\left(\mathcal{L}^{m,n}\Delta t\right) \exp\left(\mathcal{D}\Delta t/2\right), \tag{6.22}$$

where the exact form of the solution operators remains to be determined.

In the scope of this work we consider the Lindblad dissipation superoperator as the most general form and note that its treatment of the coherence terms is completely decoupled. In fact, after combining Eqs. (3.10) and (3.12) to

$$\partial_t \rho_{ij,i\neq j} = -\gamma_{ij,\mathrm{deph}} \rho_{ij}, \tag{6.23}$$

we can readily derive an elementwise update step for the coherence terms

$$\rho_{ij}^{m,n+1/2} = \exp\left(-\gamma_{ij,\mathrm{deph}}\Delta t/2\right) \rho_{ij}^{m,n-1/2}. \tag{6.24}$$

The population terms can be updated by solving Eq. (3.9) in matrix-vector form, which reads

$$\partial_t \begin{bmatrix} \rho_{11} \\ \rho_{22} \\ \vdots \\ \rho_{NN} \end{bmatrix} = \begin{bmatrix} -\tau_1^{-1} & \gamma_{12} & \cdots & \gamma_{1N} \\ \gamma_{21} & -\tau_2^{-1} & \cdots & \gamma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{N1} & \gamma_{N2} & \cdots & -\tau_N^{-1} \end{bmatrix} \begin{bmatrix} \rho_{11} \\ \rho_{22} \\ \vdots \\ \rho_{NN} \end{bmatrix} =: G\,\mathrm{diag}\left(\hat{\rho}\right). \tag{6.25}$$

Here, $\mathrm{diag}(\hat{\rho})$ denotes the vector representation of the main diagonal of the density matrix. The solution can be readily derived as

$$\mathrm{diag}\left(\hat{\rho}\right)^{m,n+1/2} = \exp\left(G\Delta t/2\right) \mathrm{diag}\left(\hat{\rho}\right)^{m,n-1/2}. \tag{6.26}$$

We note that the operations in Eqs. (6.24) and (6.26) can be applied in arbitrary order as they are completely decoupled. Also, since the dissipation operator is assumed to be time-independent, the operations can be precalculated to increase the computational efficiency.

With regard to the non-dissipative part, we can recall from Eq. (2.66) that the solution to the Liouville-von Neumann equation yields

$$\exp\left(\mathcal{L}^{m,n}\Delta t\right)\hat{\rho} = \exp\left(-\mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t\right)\hat{\rho}\exp\left(\mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t\right). \tag{6.27}$$

Naturally, this part must be calculated at every time step due to the time dependence of the electric field. The costly matrix exponential operation can now be approximated with the Cayley transform

$$\exp\left(-\mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t\right) \approx \left(\hat{I} + \mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t/2\right)^{-1} \left(\hat{I} - \mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t/2\right) \tag{6.28}$$

of the matrix $-\mathrm{i}\hbar^{-1}\hat{H}^{m,n}\Delta t$, as already described by Eq. (6.13). As possible next step, Fadeev formulae are provided in [105]) that aim to optimize the calculation of the Cayley transform. While the expressions for $N \leq 4$ are reasonably compact, no general expression for $N > 4$ is given. However, seeing that modern libraries for matrix operations are able to select suitable algorithms for certain problems automatically, we defer the derivation of a general expression and leave the optimization of the expression above in the responsibility of the underlying library.

In the mbsolve project, the update equations (6.24), (6.26), (6.27) and (6.28) have been bundled together with Eq. (6.8) to form the **os-reg-cayley** algorithm for the master equation.

6.3.2 Operator splitting in coherence vector representation (os-cvr-rodr)

The alternative method discussed in this subsection bases on a different representation of the density matrix. Whereas in the regular case the density operator is a matrix, representations as vector are also possible. We have already seen the vectorized version $\boldsymbol{\rho} = \mathrm{col}(\hat{\rho})$ in Liouville space and note that most of the elements in this vector are complex and redundant due to the properties of the density matrix. By using traceless Hermitian basis operators $\hat{F}_k$ and the Hilbert-Schmidt inner product $\langle \hat{A} | \hat{B} \rangle = \mathrm{Tr}\{\hat{A}^\dagger \hat{B}\}$, the density matrix can be transformed to a real-valued, non-redundant vector. The operators $\hat{F}_k$ from Section 3.1.2 are a reasonable choice as they are traceless and constitute an orthonormal basis. It should be noted, however, that the original $N(N-1)$ off-diagonal operators must be combined linearly to yield the Hermitian operator pairs

$$\begin{aligned} \hat{F}_{k'} &= |i\rangle \langle j| + |j\rangle \langle i|, \\ \hat{F}_{k''} &= -\mathrm{i}\,|i\rangle \langle j| + \mathrm{i}\,|j\rangle \langle i|, \end{aligned} \tag{6.29}$$

where the indices i and j are mapped suitably to $N \leq k' < N(N+1)/2$ and $N(N+1)/2 \leq k'' \leq N^2 - 1$. We project the density matrix onto the basis $\hat{F}_k$ using

$$\hat{\rho} = \frac{1}{N}\hat{I} + \frac{1}{2}\sum_{k=1}^{N^2-1} \langle \hat{F}_k | \hat{\rho} \rangle \, \hat{F}_k = \frac{1}{N}\hat{I} + \frac{1}{2}\sum_{k=1}^{N^2-1} S_k \hat{F}_k, \tag{6.30}$$

where the identity matrix $\hat{I}$ is required to satisfy the trace condition of the density matrix as the operators $\hat{F}_k$ are traceless. The resulting vector $\boldsymbol{S}$ with the coefficients $S_k = \langle \hat{F}_k | \hat{\rho} \rangle$ is called Bloch vector for $N = 2$ and referred to as coherence vector in the general case [90].

The master equation (3.69) can be transformed similarly and reads (elementwise)

$$\begin{aligned} \langle \hat{F}_k | \partial_t \hat{\rho} \rangle &= \partial_t S_k = \\ \langle \hat{F}_k | \mathcal{L}(\hat{\rho}) + \mathcal{D}(\hat{\rho}) \rangle &= \sum_{j=1}^{N^2-1} \frac{1}{2} \langle \hat{F}_k | \mathcal{L}(\hat{F}_j) \rangle \, S_j + \sum_{j=1}^{N^2-1} \frac{1}{2} \langle \hat{F}_k | \mathcal{D}(\hat{F}_j) \rangle \, S_j + \frac{1}{N} \langle \hat{F}_k | \mathcal{D}(\hat{I}) \rangle, \end{aligned} \tag{6.31}$$

since the identity matrix is time-independent, $\langle \hat{F}_i | \hat{F}_j \rangle = 2\delta_{ij}$ holds, and the action $\mathcal{L}(\hat{I})$ of the Liouvillian on the identity matrix is zero. The elementwise form can be transferred to matrix-vector notation, in which the master equation reads

$$\partial_t \boldsymbol{S} = (L + D)\,\boldsymbol{S} + \boldsymbol{S}^{\mathrm{eq}}, \tag{6.32}$$

where the elements of the $(N^2 - 1) \times (N^2 - 1)$ real matrices L and D read $L_{jk} = \langle \hat{F}_k | \mathcal{L}(\hat{F}_j) \rangle / 2$ and $D_{jk} = \langle \hat{F}_k | \mathcal{D}(\hat{F}_j) \rangle / 2$, respectively, and the $N^2 - 1$ elements of the real vector $\boldsymbol{S}^{\mathrm{eq}}$ read $S_k^{\mathrm{eq}} = \langle \hat{F}_k | \mathcal{D}(\hat{I}) \rangle / N$.

For a similar treatment of the polarization term in Eq. (6.8) we need a description of the dipole moment operator $\hat{\mu}_z$ in the coherence vector representation. This can be achieved with the projection

$$\hat{\mu}_z = \frac{1}{N} \langle \hat{\mu}_z | \hat{I} \rangle \, \hat{I} + \frac{1}{2}\sum_{k=1}^{N^2-1} \langle \hat{F}_k | \hat{\mu}_z \rangle \, \hat{F}_k, \tag{6.33}$$

where we use the identity operator as $\hat{F}_{N^2}$ (cf. Section 3.1.2). Then, we can address the essential term

$$\mathrm{Tr}\left\{\hat{\mu}_z \left[\mathcal{L}_0(\hat{\rho}) + \mathcal{D}(\hat{\rho})\right]\right\} = \frac{1}{N} \langle \hat{\mu}_z | \hat{I} \rangle \underbrace{\mathrm{Tr}\left\{\hat{I}\left[\mathcal{L}_0(\hat{\rho}) + \mathcal{D}(\hat{\rho})\right]\right\}}_{=0} + \frac{1}{2}\sum_{k=1}^{N^2-1} \langle \hat{F}_k | \hat{\mu}_z \rangle \, \langle \hat{F}_k | \mathcal{L}_0(\hat{\rho}) + \mathcal{D}(\hat{\rho}) \rangle \tag{6.34}$$

of Eq. (6.8), where Liouvillian $\mathcal{L}_0$ considers only the Hamiltonian $\hat{H}_0$ and ignores the interaction term $\hat{\mu}_z E_z$, and simplify it using the relation $\mathrm{Tr}\{\hat{a}\hat{b}\} = 0$ for a diagonal operator $\hat{a}$ and a traceless operator $\hat{b}$. Then, we can plug in the remainder into Eq. (6.8), which yields

$$\begin{aligned}(\partial_t P_z)^{m,n+1/2} &= n_{3\mathrm{D}} \frac{1}{2} \sum_{k=1}^{N^2-1} \langle \hat{F}_k | \hat{\mu}_z \rangle \langle \hat{F}_k | \mathcal{L}_0 (\hat{\rho}) + \mathcal{D} (\hat{\rho}) \rangle \\ &= \frac{1}{2} n_{3\mathrm{D}} \boldsymbol{m}^{\mathrm{T}} \left[(L_0 + D) \boldsymbol{S} + \boldsymbol{S}^{\mathrm{eq}} \right],\end{aligned} \tag{6.35}$$

where $m_k = \langle \hat{F}_k | \hat{\mu}_z \rangle$ denote the $N^2 - 1$ elements of the vector $\boldsymbol{m}$, and the matrix L_0 contains the elements $L_{0,jk} = \langle \hat{F}_k | \mathcal{L}_0 (\hat{F}_j) \rangle / 2$.

Equation (6.32) can now be treated with an operator splitting approach that considers the solutions to the equations

$$\partial_t \boldsymbol{S} = (L_0 + D) \boldsymbol{S} + \boldsymbol{S}^{\mathrm{eq}}, \tag{6.36a}$$

$$\partial_t \boldsymbol{S} = L_1 \boldsymbol{S}, \tag{6.36b}$$

where L_0 and L_1 consider the contribution of the time-independent Hamiltonian $\hat{H}_0$ and the time-dependent interaction term $\hat{\mu}_z E_z$, respectively. Since $L_0 + D$ is time-independent, the solution of Eq. (6.36a) can be determined as

$$\boldsymbol{S} = \exp \left[(L_0 + D) (t - t_0) \right] \left[\boldsymbol{S}(t_0) + (L_0 + D)^{-1} \boldsymbol{S}^{\mathrm{eq}} \right] - (L_0 + D)^{-1} \boldsymbol{S}^{\mathrm{eq}}, \tag{6.37}$$

where the inhomogeneous part of the solution $\boldsymbol{S}^{\mathrm{in}} = (L_0 + D)^{-1} \boldsymbol{S}^{\mathrm{eq}}$ is introduced for brevity. Similarly, the solution

$$\boldsymbol{S} = \exp \left[\int_{t_0}^{t} L_1(\tau) \mathrm{d}\tau \right] \boldsymbol{S}(t_0) \tag{6.38}$$

of Eq. (6.36b) can be derived, where we allow for a time-dependent L_1. After inspection of the elements

$$L_{1,jk} = \frac{1}{2} \langle \hat{F}_k | - \mathrm{i}\hbar^{-1} \left[-\hat{\mu}_z E_z, \hat{F}_j \right] \rangle \coloneqq M_{jk} E_z \tag{6.39}$$

of L_1 it becomes apparent that they are proportional to the electric field and that the newly introduced matrix M is time-independent. Again, we follow the recommendation in [106] and solve the dissipative, time-independent part for a half time step, the non-dissipative, time-dependent part for a full step, and repeat the first step. This leads to the update equations

$$\boldsymbol{S}' = \exp \left[(L_0 + D) \Delta t / 2 \right] \left(\boldsymbol{S}^{m,n-1/2} + \boldsymbol{S}^{\mathrm{in}} \right) - \boldsymbol{S}^{\mathrm{in}}, \tag{6.40a}$$

$$\boldsymbol{S}'' = \exp \left[M E_z^{m,n} \Delta t \right] \boldsymbol{S}', \tag{6.40b}$$

$$\boldsymbol{S}^{m,n+1/2} = \exp \left[(L_0 + D) \Delta t / 2 \right] \left(\boldsymbol{S}'' + \boldsymbol{S}^{\mathrm{in}} \right) - \boldsymbol{S}^{\mathrm{in}}, \tag{6.40c}$$

where the integral expression

$$\int_{(n-1/2)\Delta t}^{(n+1/2)\Delta t} L_1(\tau) \mathrm{d}\tau \approx M E_z(n\Delta t) \Delta t \tag{6.41}$$

has been approximated using the midpoint rule. Of course, we would have achieved the same result by assuming that the electric field does not vary during the update of the density matrix, as we have done with the other weakly coupled methods. As we can calculate the inhomogeneous vector $\boldsymbol{S}^{\mathrm{in}}$ and the matrix exponential $\exp \left[(L_0 + D) \Delta t / 2 \right]$ once and reuse it every time step, we focus in the following on the efficient calculation of the exponential $\exp \left[M E_z^{m,n} \Delta t \right]$.

Due to the cyclic property of the trace operation, we can derive an alternative expression for the elements

$$M_{jk} = \frac{1}{2} \mathrm{i}\hbar^{-1} \langle \hat{F}_k | \left[\mu_z, \hat{F}_j \right] \rangle = \frac{1}{2} \hbar^{-1} \langle \hat{\mu}_z | \mathrm{i} \left[\hat{F}_j, \hat{F}_k \right] \rangle =: \frac{1}{2} \hbar^{-1} \langle \hat{\mu}_z | \hat{C} \rangle \tag{6.42}$$

of the matrix M. We can readily see that

$$\hat{C}^\dagger = \left(\mathrm{i}\left[\hat{F}_j, \hat{F}_k\right]\right)^\dagger = -\mathrm{i}\hat{F}_k\hat{F}_j + \mathrm{i}\hat{F}_j\hat{F}_k = \mathrm{i}\left[\hat{F}_j, \hat{F}_k\right] = \hat{C} \tag{6.43}$$

holds and the commutator term $\hat{C}$ is Hermitian. Here, the Hermitian property of the basis operators $\hat{F}_k$ is used. Then, using again the properties of the trace operation and the Hermitian property of $\hat{\mu}_z$, we can verify that

$$\langle\hat{\mu}_z|\hat{C}\rangle = \mathrm{Tr}\left\{\hat{\mu}_z\hat{C}\right\} = \mathrm{Tr}\left\{\hat{C}^\mathrm{T}\hat{\mu}_z^\mathrm{T}\right\} = \mathrm{Tr}\left\{\hat{\mu}_z^\mathrm{T}\hat{C}^\mathrm{T}\right\} = \mathrm{Tr}\left\{\hat{\mu}_z^*\hat{C}^*\right\} = \mathrm{Tr}\left\{\left(\hat{\mu}_z\hat{C}\right)^*\right\} = \mathrm{Tr}\left\{\hat{\mu}_z\hat{C}\right\}^* \tag{6.44}$$

holds and deduce that the elements M_{jk} are real. Finally, due to the properties of the commutator we can write

$$M_{kj} = \frac{1}{2}\mathrm{i}\hbar^{-1}\langle\hat{\mu}_z|\mathrm{i}\left[\hat{F}_k, \hat{F}_j\right]\rangle = -\frac{1}{2}\mathrm{i}\hbar^{-1}\langle\hat{\mu}_z|\mathrm{i}\left[\hat{F}_j, \hat{F}_k\right]\rangle = -M_{jk} \tag{6.45}$$

and note that the elements M_{jk} are antisymmetric.

For the exponential of antisymmetric matrices there exist specialized formulas and algorithms. The special case $N = 2$ yields a 3×3 matrix

$$M = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}, \tag{6.46}$$

and the corresponding matrix exponential can be calculated using Rodrigues' formula

$$\exp\left(ME_z^{m,n}\Delta t\right) = I + \frac{\sin\left(\theta E_z^{m,n}\Delta t\right)}{\theta}M + \frac{1 - \cos\left(\theta E_z^{m,n}\Delta t\right)}{\theta^2}M^2, \tag{6.47}$$

where I is the 3×3 identity matrix and $\theta = \sqrt{a^2 + b^2 + c^2}$. This formula can be generalized to arbitrary matrix dimensions [168] and then assumes the form

$$\exp\left(ME_z^{m,n}\Delta t\right) = I + \sum_{i=1}^{p}\sin\left(\theta_i E_z^{m,n}\Delta t\right)B_i + \left[1 - \cos\left(\theta_i E_z^{m,n}\Delta t\right)\right]B_i^2, \tag{6.48}$$

where I is now the $N^2 - 1 \times N^2 - 1$ identity matrix and p is the number of distinct, non-zero eigenvalue pairs $\theta_{2i-1,2i} = \pm\mathrm{i}\theta_i$ of the matrix M. For each eigenvalue pair there exists a corresponding, antisymmetric matrix

$$B_i = T\Theta_i T^\mathrm{T} \tag{6.49}$$

that is composed of an orthogonal matrix T and the matrix Θ_i. The same matrix T appears in a decomposition $M = T\Theta T^\mathrm{T}$, where Θ is a block-diagonal matrix with a block

$$\theta_i \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{6.50}$$

for each eigenvalue pair[168], [169]. The matrix Θ_i, on the other hand, consists of a single, normalized block

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{6.51}$$

in the $(2i - 1)$-th and $(2i)$-th row and column, respectively, and is zero otherwise. Thereby, the matrix exponential can be reduced to calls to trigonometric functions, matrix scaling, and matrix addition operations. This assumes, of course, that the eigenvalues θ_i and the matrices B_i and B_i^2 are precalculated, where for the latter the relation $B_i^2 = T\Theta_i^2 T^\mathrm{T}$ can be used.

In the mbsolve project, the update equations (6.40) and (6.48) have been bundled together with Eq. (6.35) to form the **os-cvr-rodr** algorithm for the master equation.

6.3.3 Action approach in real-valued vector representation (action-rvr)

Vectorized representations of the density matrix, such as $\boldsymbol{\rho}$ in Liouville space or the coherence vector $\boldsymbol{S}$, allow mathematically compact solutions. The major drawback of these approaches is that the involved matrices grow exponentially with the number of levels N and typically have the dimensions $N^2 \times N^2$ (in Liouville space) or $(N^2-1) \times (N^2-1)$ (in coherence vector representation). As a consequence, certain operations such as the matrix-vector multiplications required in Eq. (6.40) do not scale well when increasing the number of levels. Nevertheless it would be tempting to use a vectorized representation that allows a general and exact solution (possibly without invoking the operator splitting technique). In this section, a promising approach is discussed.

A representation similar to the coherence vector can be written as

$$\hat{\rho} = \sum_{k=1}^{N^2} \langle \hat{G}_k | \hat{\rho} \rangle \, \hat{G}_k = \sum_{k=1}^{N^2} R_k \hat{G}_k, \tag{6.52}$$

where the N^2 elements R_k form the real-valued vector $\boldsymbol{R}$. The basis operators G_k consist of N diagonal operators

$$\hat{G}_k = |k\rangle \langle k|, \tag{6.53}$$

where $k \in [1, N]$, and $(N-1)N$ off-diagonal operator pairs

$$\begin{aligned} \hat{G}_{k'} &= \frac{1}{\sqrt{2}} \left(|i\rangle \langle j| + |j\rangle \langle i| \right), \\ \hat{G}_{k''} &= \frac{-\mathrm{i}}{\sqrt{2}} \left(|i\rangle \langle j| - |j\rangle \langle i| \right), \end{aligned} \tag{6.54}$$

where the indices i and j are mapped suitably to $N < k' \leq N(N+1)/2$ and $N(N+1)/2 < k'' \leq N^2$. In contrast to the coherence vector representation, the square root in the off-diagonal operators assures that $\langle \hat{G}_i | \hat{G}_j \rangle = \delta_{ij}$ and the operators are orthonormal. We can readily see that the operators form a basis of all Hermitian matrices (as opposed to the basis $\hat{F}_k$ of traceless Hermitian matrices), hence no extra term involving the identity matrix is required in Eq. (6.52). In the following, we refer to this representation as real-valued vector representation due to the lack of an established name. In this representation, the master equation (3.69) reads

$$\partial_t \boldsymbol{R} = (L + D) \, \boldsymbol{R}, \tag{6.55}$$

where L and D are real $N^2 \times N^2$ matrices with the elements $L_{jk} = \langle \hat{G}_k | \mathcal{L}(\hat{G}_j) \rangle$ and $D_{jk} = \langle \hat{G}_k | \mathcal{D}(\hat{G}_j) \rangle$, respectively. For Eq. (6.8) we can write

$$(\partial_t P_z)^{m,n+1/2} = n_{3\mathrm{D}} \sum_{k=1}^{N^2} \langle \hat{G}_k | \hat{\mu}_z \rangle \, \langle \hat{G}_k | \mathcal{L}_0 (\hat{\rho}) + \mathcal{D} (\hat{\rho}) \rangle = n_{3\mathrm{D}} \boldsymbol{m}^{\mathrm{T}} (L_0 + D) \, \boldsymbol{R}, \tag{6.56}$$

where in this case $m_k = \langle \hat{G}_k | \hat{\mu}_z \rangle$ denote the N^2 elements of the vector $\boldsymbol{m}$, and the matrix L_0 contains the elements $L_{0,jk} = \langle \hat{G}_k | \mathcal{L}_0 (\hat{G}_j) \rangle$.

Equation (6.55) yields the solution

$$\boldsymbol{R}^{m,n+1/2} = \exp\left[(L + D) \, \Delta t \right] \boldsymbol{R}^{m,n-1/2} \tag{6.57}$$

for a single time step, where the electric field is incorporated in the matrix L and again assumed to be constant during the update of $\boldsymbol{R}$. We note that explicit knowledge of the matrix exponential is not required but only its action on the vector $\boldsymbol{R}$. For this range of applications a promising algorithm exists [153]. It bases on the scaling and squaring approach [152]

$$\exp(A \Delta t) = \prod^{s} \exp\left(s^{-1} A \Delta t \right) \approx \prod^{s} r_m \left[\exp\left(s^{-1} A \Delta t \right) \right], \tag{6.58}$$

where the parameter s is used to scale the matrix A, so that the resulting matrix exponential $\exp(s^{-1}A\Delta t)$ lies in a regime in which it can be approximated by r_m with negligible numerical error. The approximation r_m, which is frequently achieved by using Padé approximants [152], consists in this case of the truncated Taylor polynomial

$$\exp\left(s^{-1}A\Delta t\right) \approx \sum_{k=0}^{m} \frac{s^{-k}A^k\Delta t^k}{k!} =: r_m\left[\exp\left(s^{-1}A\Delta t\right)\right]. \tag{6.59}$$

Then, the s squaring steps required to calculate the action

$$\exp\left(A\Delta t\right)\boldsymbol{b} \approx \prod^{s} r_m\left[\exp\left(s^{-1}A\Delta t\right)\right]\boldsymbol{b} = (r_m(\cdots(r_m(r_m\boldsymbol{b}))))) \tag{6.60}$$

of the matrix exponential on the vector $\boldsymbol{b}$ can be expressed as iterative updates

$$\boldsymbol{b} \leftarrow r_m\boldsymbol{b} = \sum_{k=0}^{m} \frac{s^{-k}A^k\Delta t^k}{k!}\boldsymbol{b} = \sum_{k=1}^{m}(s_k(\cdots(s_2(s_1(\boldsymbol{b})))))+\boldsymbol{b}, \tag{6.61}$$

which in turn use the iterative updates

$$\boldsymbol{b} \leftarrow s_j\boldsymbol{b} = \frac{\Delta t}{js}A\boldsymbol{b}. \tag{6.62}$$

Since the update chains $s_k(\cdots(s_1\boldsymbol{b}))$ can be reused efficiently in the sum, the cost of one density matrix update is dominated by ms matrix-vector products. Also, the accuracy of the algorithm is completely controlled by the parameters m and s. Fortunately, procedures are described that determine the optimal parameter choice for a given matrix A [153]. Thereby, the optimal tradeoff between computational effort and accuracy is achieved. In a practical realization, the density vector $\boldsymbol{R}$ can be updated using one auxiliary vector. Apart from that, only the matrix $A = (L + D)$ needs to be stored. Additionally, the matrix A is likely sparse for larger N.

In the mbsolve project, the update equations (6.57) and the algorithm for the action of the matrix exponential [153] have been bundled together with Eq. (6.56) to form the **action-rvr** algorithm for the master equation.

6.3.4 Single-thread performance measurement

The performance of the three algorithms for the master equations is now evaluated. As test example, we use once more the anharmonic ladder example from Section 4.3.3 [98], where we vary the number of energy levels N between 2 and 10. In principle an example with any spatial dimension would not be required for the discussion of numerical methods for the master equation. However, it makes sense to use such an example nevertheless, as we discuss methods for the generalized Maxwell-Bloch equations and e.g., the polarization term calculation could affect the performance as well. It should be noted that all three algorithms are weakly coupled to the same FDTD implementation.

The algorithms are implemented in the scope of the mbsolve project and compiled with the GNU C++ compiler 7.3.0 and the compiler flags `-O3 -march=native`. For the matrix operation, the Eigen library 3.3.4 is used [151]. Then, the test example is simulated on a single Intel Xeon E5-2697 v3 processor core using a single thread. Thereby, parallelization is disabled and does not affect the performance measurements. As main figure of merit the metric grid point updates per second (GPU/s) is used, which is calculated as

$$P = \frac{N_x N_t}{t_{\text{exec}}}, \tag{6.63}$$

where N_x and N_t denote the number of spatial and temporal grid points, respectively, and t_{exec} is the measured execution time. This allows the comparison of e.g., simulations with different simulation end time, which is typically adjusted to ensure that the simulation runs for at least 30 s. Otherwise, artifacts from e.g., the underlying operating system could affect the measurements. In order to eliminate other possible fluctuations, the measurements are repeated five times and the time measurements are averaged [109]–[111].

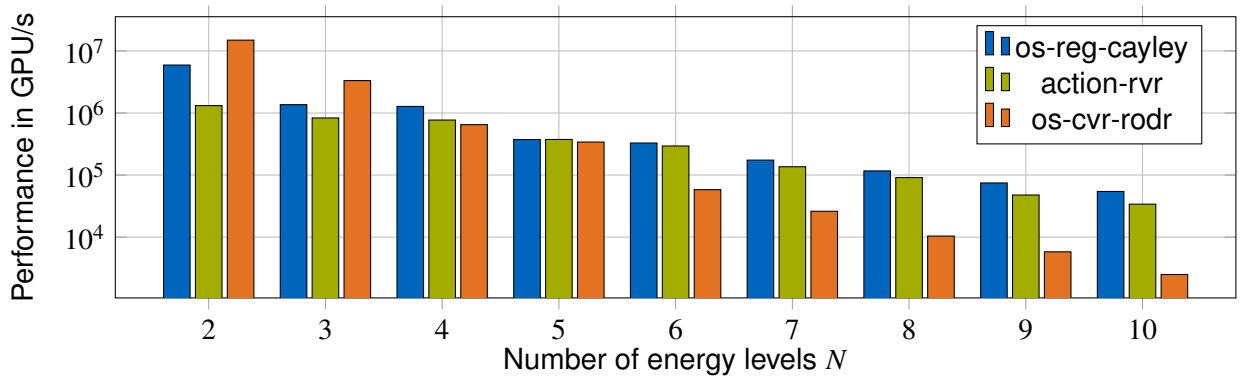


Figure 6.3 Single-thread performance of the three methods that solve the master equation numerically. Reprinted with permission from M. Riesch *et al.*, "Completely positive trace preserving methods for the Lindblad equation" [111]. © 2020 IEEE.

The performance results for the three algorithms are depicted in Fig. 6.3 [111]. Compared to the reference algorithm **os-reg-cayley**, the performance of the **os-cvr-rodr** approach is larger for large N but then drops significantly with increasing N. Seeing that simulations with only two or three energy levels are very common, this algorithm provides an interesting alternative as no additional error due to the Cayley approximation is introduced. The **action-rvr** algorithm features the least numerical error, as neither the operator splitting technique nor other approximations are invoked. Additionally, its performance scales quite robustly with the number of energy levels N. It should be noted that the optimal parameters m and s were not determined at every update of the density matrix, as the naive implementation of the procedures in [153] lead to inferior performance. It has been found, however, that the parameters remain the same over the complete simulation and lead to the correct simulation results. Therefore, the performance results can be considered as best-case scenario. They show that the **action-rvr** remains to be an interesting alternative with superior accuracy and reasonable performance, given that the parameter determination procedures are optimized in future work.

6.4 Summary

The rotating wave approximation (RWA) is frequently used in related work, but has been dismissed as unsuitable for the simulation of frequency comb operation or ultra-short pulse generation scenarios. Therefore, numerical methods that solve Maxwell's equation directly (full-wave simulation) have been discussed. From the two candidates used in related work the finite-difference time-domain (FDTD) method has been selected as basis for the next steps, as it can be implemented with relatively little effort and provides reasonable performance. As next step, the evaluation of methods that consider an adaptive grid would be promising, as the FDTD is bound to a constant spatial and temporal discretization size.

In order to couple Maxwell's equations to the master equation, we have defined an abstract update operator of the density matrix as placeholder, and have discussed possible discretizations of the density matrix. We have found that the weakly coupled scheme is well-suited for the integration of the update operator into the spatiotemporal grid of the FDTD. The update operator must preserve the properties of the density matrix, especially in long-term simulations. This is the case if the update operator is a completely positive trace-preserving (CPTP) map. We have reviewed the numerical methods for the master equation in related work, and have discussed in detail which methods feature a CPTP update operator.

From the wide range of candidates, a well-established reference method as well as two promising alternatives have been selected. The three methods have been presented in detail, where the implementation in the scope of the mbsolve project has been in the focus. The performance of the three methods has been measured for a simulation problem, in which the numbers of energy levels varies from 2 to 10. We have found that while the reference method displays the best overall performance, it is outperformed by the first alternative for the special, but crucial case of two energy levels. Furthermore, we have discussed the

potential of the second alternative, which features the largest accuracy and reasonable performance. Here, it should be noted that future work needs to revise the calculation of the parameters in order to provide an optimal tradeoff between performance and accuracy. Additionally, further optimizations of the numerical methods with respect to static or dynamic allocation of matrices and vectors (particularly relevant for large, sparse matrices) and the exploitation of the Hermitian properties of certain matrices can be envisaged.

7 Solving the Maxwell-Bloch equations on parallel architectures

Nigel: What we do is, if we need that extra push over the cliff, you know what we do?
Marty: Put it up to eleven.
Nigel: Eleven. Exactly. One louder.
Marty: Why don't you just make ten louder and make ten be the top number and make that a little louder?
Nigel: [pause] These go to eleven.
— Nigel Tufnel and Marty DiBergi, This is Spinal Tap

Tim: What do we need?
Audience: More Power!
Tim: Darn right, more power!
— Tim "Toolman" Taylor, Home Improvement

Numerical simulations of QCLs cause significant computational effort due to several reasons. As we discussed in Chapter 6, the rotating wave approximation (RWA) should not be used for the simulation of frequency comb operation or ultra-short pulse generation scenarios. Instead, full-wave simulations should be carried out, which typically require a finer discretization and hence increase the computational effort. Also, often long-term simulations are required to provide the necessary spectral resolution. Furthermore, the simulations often need to consider several energy levels to model the QCL dynamics appropriately, which increases the efforts for each density matrix update significantly. As a remedy, we can divide the computational workload among many compute units in modern parallel architectures. While parallel implementations of numerical solvers for the Maxwell-Bloch equations have been presented in related literature [108], [114], [170], an assessment of the parallel efficiency has not been published to the best of our knowledge. Furthermore, most of the related projects are not open-source.

Therefore, the aim of this chapter is to provide a parallel implementation of a solver in the scope of the mbsolve project. To this end, the finite-difference time-domain (FDTD) method for Maxwell's equations is implemented using the C++ programming language and the OpenMP standard for parallelization. In this implementation, described in Section 7.1, a template mechanism allows to couple different approaches for the master equation to the FDTD method. The resulting source code is of course publicly available. However, early performance measurement results showed that this approach is not optimal for all problems [109]. As a consequence, Section 7.2 presents an advanced version of the algorithm. The parallel efficiency and performance of both implementations on a shared memory many-core system are subsequently measured in Section 7.3. Finally, we conclude with a brief summary and discuss possible extensions of the parallel implementation.

This chapter is largely based on M. Riesch, N. Tchipev, H.-J. Bungartz, *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures", in *Proceedings of the Platform for Advanced Scientific Computing Conference*, Zurich, Switzerland: ACM, New York, NY, Jun. 2019, 5:1–5:8. DOI: 10.1145/3324989.3325715, an open access article published under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

7.1 Parallel implementation of the FDTD method

As pointed out in Chapter 6 (cf. Fig. 6.1), the update of the density matrix does not have any spatial dependence. We can therefore restrict the discussion to the parallelization of the FDTD method exclusively, and include the density matrix at a later stage. The FDTD method consists in essence of the Eqs. (6.1) and (6.2). We recall that these equations are partially in implicit form and begin by rearranging them, which yields the update equations

$$\begin{aligned} E_z^{m,n+1} =& a' E_z^{m,n} - b' \Gamma \left(\partial_t P_z\right)^{m,n+1/2} + \\ & b' \Delta x_{\text{inv}} \left(H_y^{m+1/2,n+1/2} - H_y^{m-1/2,n+1/2} \right) \end{aligned} \tag{7.1}$$

with the coefficients

$$a' = \frac{1 - \Delta t (2\epsilon)^{-1} \sigma}{1 + \Delta t (2\epsilon)^{-1} \sigma}, \qquad b' = \frac{\Delta t \epsilon^{-1}}{1 + \Delta t (2\epsilon)^{-1} \sigma}, \qquad \Delta x_{\text{inv}} = \frac{1}{\Delta x},$$

and

$$H_y^{m+1/2,n+1/2} = H_y^{m+1/2,n-1/2} + c' \left(E_z^{m+1,n} - E_z^{m,n} \right) \tag{7.2}$$

with the coefficient

$$c' = \frac{\Delta t}{\Delta x \mu}.$$

The coefficients are constant (at least, constant for each material) and are precalculated once to save costly division operations. Prior to this step, the number of spatial and temporal grid points has to be determined. The user is responsible for the specification of the number of spatial grid points N_x. The discretization size $\Delta x = L/(N_x - 1)$ is calculated for a given total length L of the simulation domain, which often corresponds to the QCL cavity length L_{cav}. Here, care must be taken to choose a reasonable number of grid points so that the resulting discretization size lies in the range $\lambda/20$ to $\lambda/200$, where λ is the smallest occurring wavelength (cf. Chapter 6). The temporal discretization size Δt is calculated with Eq. (6.4), where the smallest refractive index in the simulation domain is selected, and the Courant number is set to $C = 1/2$, which was found to be adequate in related literature [32]. Then, based on the simulation end time t_{e} the number of temporal grid points can be calculated, where the discretization size Δt may be decreased slightly to allow an integer number N_t. We note that for the borderline case $N_x = 1$ the spatial discretization size calculation does not make sense. In this case, the user can specify the number of temporal grid points, which determines the temporal discretization size. The initialization of the discretization sizes, numbers of grid points, and coefficients for the update equations is delegated to the helper functions **init_fdtd_simulation**, which analyses the given device and scenario and calculates appropriate numbers of grid points, and **get_fdtd_constants**, respectively.

With regard to the update equations (7.1) and (7.2), we note that all electric and all magnetic field points can be updated simultaneously, but synchronizations are required after all electric field updates and, respectively, after all magnetic field and density matrix updates. Furthermore, the density matrix updates do not depend on their neighbors and can be processed in parallel, as we have already pointed out. Algorithm 1 shows the required calculations in the main loop in a simplified manner. Here, the abstract functions **update_h** and **update_e** represent the update equations (7.1) and (7.2), whereas **update_d** and **calc_p** refer to the update of the density matrix and the calculation of the polarization term in Eq. (6.8), which are both provided by the specified algorithm for the master equation. As their names suggest, the functions **sync** and **record_results** represent the code that ensures the synchronization of all compute units and store the desired simulations results, respectively.

By adding OpenMP directives (e.g., `#pragma omp parallel for`) to the serial code, both loops over m in Algorithm 1 can be distributed over multiple threads. Thereby, a parallelized version is created with minimal changes compared to the serial version. The details such as thread creation or synchronization are implicitly added by the compiler and in principle the user does not have to worry about the details. When it comes to the optimization of the parallel code, however, one must know the details behind the scenes. For

Algorithm 1 Simulation main loop – basic version.

```
for n = 0 to n_max do
  for m = 1 to m_max do
    h[m] ← update_h(e[m], e[m − 1])
    d[m] ← update_d(e[m])
    p[m] ← calc_p(d[m])
  end for
  sync()
  for m = 0 to m_max − 1 do
    e[m] ← update_e(h[m + 1], h[m], p[m])
  end for
  sync()
  record_results()
end for
```

example, it is feasible to create the threads before the main loop with `#pragma omp parallel` and reuse them at every loop with `#pragma omp for`. The partitioning of the data must be considered thoroughly. First, the workload per thread should be approximately equal. Also, the partition boundaries should be aligned to the size of the cache line size in order to prevent false cache sharing. And finally, the assignment of threads to processors is crucial. Although the memory appears uniform to the programmer, it consists physically of several domains. While all processors are able to access memory addresses of all domains, the latency differs significantly. Also, situations where all data are stored in one single domain must be avoided, since then an artificial bottleneck would be created. Such first-touch issues can be avoided by initializing the data in parallel and pinning the threads to the processors (and therefore to the memory domains).

In the mbsolve project, Algorithm 1 is implemented in the form of the template class **solver_cpu_fdtd**. We recall that different algorithms for the master equations (and variations thereof) are likely to be implemented and evaluated in mbsolve. Since it is considered bad practice to copy and paste parts of the code, such as the FDTD implementation, we need a way to factor out the common parts. The template class provides an elegant way to do so, since we easily can create e.g., a solver object **solver_cpu_fdtd_cvr_rodr** from the template class by providing the corresponding class **algo_lindblad_cvr_rodr** as template argument. The class to be passed as argument must provide certain methods, which are subsequently called by the template class. Most importantly, the class method **update** is called, which combines the functions **update_d** and **calc_p** in Algorithm 1. The template argument is considered at compile time, which increases the time required to compile to project significantly but offers an efficient approach at runtime.

7.2 Advanced FDTD implementation

Despite all efforts to optimize the basic version of the FDTD implementation, it turned out that in certain cases the performance did not scale very well with the number of used processors [109]. The reason for this behavior are the expensive synchronizations, which are performed twice during each iteration. If the amount of work per processor is relatively small, the synchronization overhead exceeds the benefit of using multiple processors in parallel. The remedy for this problem was a communication-reducing approach as presented in [171]. Here, each thread calculated certain field values redundantly around the partition boundary. Hence, less synchronizations were required and the parallel efficiency increased.

Algorithm 2 shows the essential elements of this approach. Here, we assume that every thread runs the main loop. First, the data in the overlap regions are exchanged between the threads. After this operation, a synchronization is required. Then, each thread performs the calculations for all grid points in its partition and in the overlap regions. Since there are no dependencies on data belonging to other threads, no synchronizations are necessary. After the sub-loop, i.e., before the next data exchange, a synchronization call must be inserted using the `#pragma omp barrier` pragma. Figure 7.1 sketches an example of this

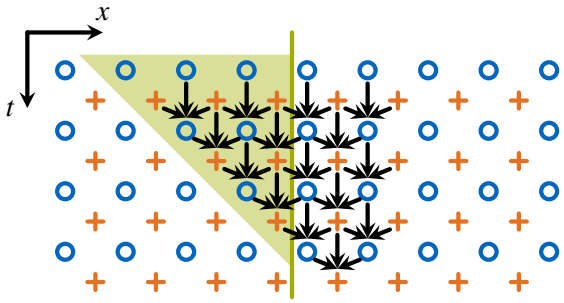


Figure 7.1 Advanced version of the FDTD with redundant calculations at the partition boundary. Instead of performing two costly synchronizations at each iteration, the data in the overlap region (size $OL = 3$) are exchanged between the threads. In order to calculate the electric field (orange crosses) at the boundary, the calculations in the green triangle are required. Those are performed by both neighbors in a redundant fashion. Therefore, four synchronization calls are traded against twelve extra calculations. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

technique. Here, at each boundary two regions with the size $OL = 3$ are exchanged. The number of redundant calculations is in this case $2 \cdot 2 \cdot 3 = 12$ (two regions, two fields, $\sum_m^{OL-1} m = OL(OL-1)/2 = 3$). On the other hand, while originally $2OL = 6$ synchronizations were required, the advanced version only needs 2.

Algorithm 2 Simulation main loop – advanced version.

```
for n = 0 to n_max/OL do
    exchange_data(OL)
    sync()
    for l = 0 to OL do
        for m = 1 − OL to partition_size + OL do
            h[m] ← update_h(e[m], e[m − 1])
            d[m] ← update_d(e[m])
            p[m] ← calc_p(d[m])
        end for
        for m = −OL to partition_size + OL − 1 do
            e[m] ← update_e(h[m + 1], h[m], p[m])
        end for
    end for
    sync()
end for
```

Since a trade-off between redundant calculations and synchronizations is performed, a tuning parameter OL, which is the size of the overlap region, is introduced. The optimal choice of OL is likely to be different for each parallel architecture and for each problem size (in particular, the number of energy levels), respectively, and has to be determined carefully by performance measurements. After knowledge of the optimal choice has been gained, the simulation software can be adjusted to select automatically a reasonable overlap region size for a given problem and architecture.

In the mbsolve project, Algorithm 2 is implemented in the form of the template class **solver_cpu_fdtd_red**, where the suffix **red** refers to the redundant calculations used in this approach. Similar to **solver_cpu_fdtd**, this template class expects a class for the master equation algorithm as template argument, where the classes to be passed are of course compatible between the two FDTD implementations.

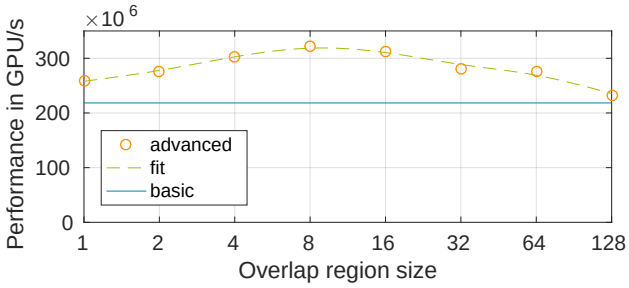


Figure 7.2 Performance of the advanced implementation (using 28 threads) for the two-level test case as function of different overlap region sizes. The performance of the basic version is shown as reference. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

7.3 Performance measurements

In this section, the performance measurement results of both versions of the FDTD implementation are shown. To this end, the source code is compiled using the GNU compiler collection (g++ 7.3.0) with the compiler optimization flags `-O3 -march=native` and the Eigen library version 3.3.4. Subsequently, exclusive access to a shared memory node with two Intel Xeon E5-2697 v3 processors (with 14 physical cores each and Intel HyperThreading disabled) is acquired. The LIKWID [172] tool suite is used to pin the threads to the processor cores as close to each other as possible. During all performance measurements, it is ensured that each measurement has a duration of at least 30 s in order to eliminate variations from the operating system. Furthermore, each measurement is repeated five times and the average performance value is taken. The parallel efficiency $E_T = S_T/T$ is used to evaluate to what degree the T threads are utilized. Here, the speedup $S_T = P_T/P_1$, where P_T is given by Eq. (6.63) and denotes the the performance with T used threads. Two setups are used to assess the performance of the implementations: a two-level QCL simulation [25] and the six-level setup presented in [98]. Both setups are discretized with 8192 as well as with 16384 spatial grid points.

7.3.1 Two-level setup

For a simple two-level simulation, the computational workload (and hence the cost for redundant calculations) is relatively small. In such cases it should be beneficial to use the advanced implementation. Therefore, we compile our code with different overlap sizes OL, which has been introduced as compile-time constant variable to allow optimizations by the compiler, and measure the performance. The measurement results in Fig. 7.2 clearly show that the advanced version outperforms the basic implementation over the chosen range of values for OL. Furthermore, the performance peak can be easily identified to be at $OL = 8$, where the advanced version provides an improvement of 45 % of the basic version's performance.

Using this overlap size, we compare the performance scalability of both implementations. The performance curves in Fig. 7.3 demonstrate that the advanced version is clearly superior for this particular simulation problem. While the basic version's performance saturates for higher number of used threads, the advanced version achieves the parallel efficiency $E_{28} = 79.4\,\%$ for 8192 grid points and $E_{28} = 84\,\%$ for 16384 grid points.

7.3.2 Six-level setup

For the six-level test case, the situation is quite different. Here, the density matrix updates are exceedingly expensive which is – at least in terms of parallel efficiency – beneficial. The results in Fig. 7.4 confirm this statement. Only the lowest possible overlap value $OL = 1$ yields reasonable performance, whereas with a larger overlap region size the advanced implementation is outperformed by the basic version.

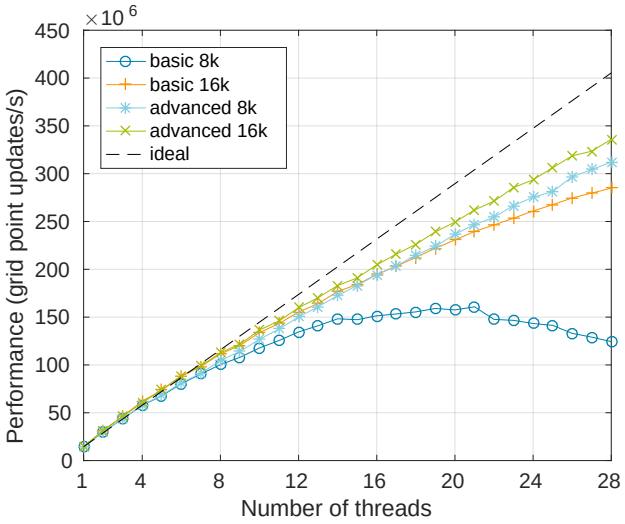


Figure 7.3 Performance scalability of both implementations for the two-level test case using 8192 and 16384 grid points, respectively. For the advanced version, $OL = 8$ was chosen. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

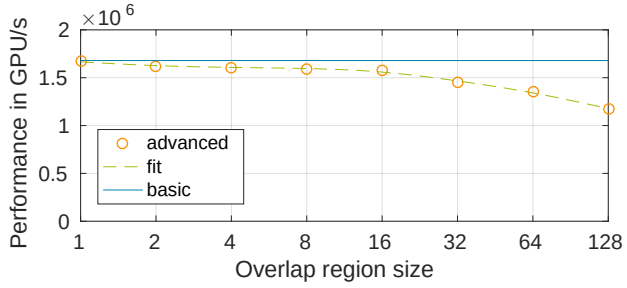


Figure 7.4 Performance of the advanced implementation (using 28 threads) for the six-level test case as function of different overlap region sizes. The performance of the basic version is shown as reference. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

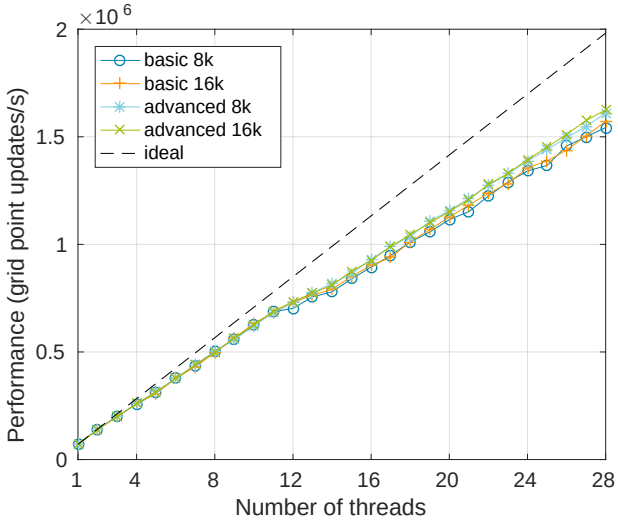


Figure 7.5 Performance scalability of both implementations for the six-level test case using 8192 and 16384 grid points, respectively. For the advanced version, $OL = 1$ was chosen. Reprinted from M. Riesch *et al.*, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures" [146] (CC BY 4.0).

Again, we haven chosen the overlap region size with the best performance and performed the scalability measurement. As Fig. 7.5 shows, the performance of both implementations are comparable, i.e., the efforts of the advanced version are not beneficial in this particular case. The parallel efficiency of the advanced implementation exceeds 82 % for both problem sizes.

7.4 Summary

We have discussed two parallelized implementations of a numerical method for the Maxwell-Bloch equations based on the finite-difference time-domain method. The straight-forward implementation using simple OpenMP directives has not shown the desired parallel efficiency in test cases with relatively low computational workload. In these cases, an advanced version that exchanges costly synchronizations against fast redundant calculations is clearly beneficial. Here, the optimal overlap region size must be determined for each combination of simulation problem and architecture.

The measured parallel efficiency values of ≈ 80 % and above confirm that the respective implementations are reasonably efficient. The deviations from the ideal curve can be explained by considering the grid size. Using 8192 grid points and 28 processor cores, each thread operates on ≈ 300 grid points, which does not yield a large amount of calculations (given the performance of recent processor cores). As a consequence, the overhead to coordinate the cores is relatively large. As the plots in Figs. 7.3 and 7.5 demonstrate, the efficiency tends to increase for larger problems, i.e., with more grid points or with more energy levels.

Those findings are quite common in high-performance computing and can be readily transferred to other parallel architectures. For example, the synchronizations are bound to become even more costly on distributed memory systems due to the larger communication costs. Also, larger problems (with respect to the spatial grid size) generally benefit more from parallelization. Therefore, we conclude that careful optimization for a certain combination of problem size, numerical method for the master equation, and the parallel architecture in use is still required. However, the parallel implementations discussed in this chapter already constitute efficient tools to run Maxwell-Bloch simulations in parallel.

8 Application of the simulation framework to different problems

> Aber immer mehr kommen wir zu dem schrecklichen Schluß: unsere Radium-Isotope verhalten sich nicht wie Radium, sondern wie Barium. Wie gesagt, andere Elemente, Trans-Urane kommen nicht in Frage. Ich habe mit Straßmann verabredet, dass wir vorerst nur Dir dies sagen wollen. Vielleicht kannst Du irgendeine phantastische Erklärung vorschlagen.
>
> — *Otto Hahn, in a letter to Lise Meitner*

At this point, we can apply the mathematical models, the numerical methods, and the mbsolve software project discussed in the previous chapters and simulate different quantum cascade laser setups. In the following, one section is dedicated to the description of each setup followed by the interpretation of the corresponding simulation results. The simulation setup in Section 8.1 reproduces the results of a theoretical study [25] and serves as additional verification example, whereas the setups in the Sections 8.2 and 8.3 aim to answer questions raised by the results of experimental measurements [20], [173]. The final section of this chapter summarizes the current applications of the work at hand and gives a brief outlook on future extensions and use cases.

8.1 Colliding pulse mode-locking in quantum cascade lasers

As already stated in Chapter 1, mode-locking of quantum cascade lasers without an external modulation (i.e., passive mode-locking, PML) has not yet been achieved. Although early studies reported successful attempts to mode-lock a QCL passively [174], [175], it was shown later that the experimental proof was not sufficient and also dynamic instabilities could be the cause of the demonstrated experimental results [176]. Furthermore, it is believed that PML in the traditional sense [16] is impossible due to the short relaxation times of QCLs [176]. Therefore, a theoretical study [25] explored alternative mode-locking schemes including the so-called colliding pulse mode-locking (CPML) approach. Simulations of this CPML setup demonstrated ultra-short pulse operation, where in contrast to traditional passive mode-locking two pulses per round trip were emitted.

In this section, we aim to reproduce the simulation of the CPML setup in [25] for two reasons: to have an additional verification example for the mbsolve solver tool (in the ultra-short pulse regime) and to discuss the effects of invoking the RWA (as was done in [25]). The CPML setup, depicted in Fig. 8.1, consists of a sequential arrangement of two gain sections, embedding an absorber section with the same transition frequency. While the gain sections and the absorber section can be realized by using different active region designs, it has been pointed out that a similar result can be achieved by using different bias voltages [77]. In [25] both gain and absorber sections are represented by two-level systems with the common transition frequency $\omega_{21} = 2\pi 3.4\,\mathrm{THz}$, where the remaining parameters may differ (cf. Tab. 1 in [25]). With the code snippet in Listing 8.1, this model can be implemented effortlessly in mbsolve.

Listing 8.1 Code snippet of the mbsolve-tool C++ application that reproduces the colliding pulse mode-locking simulation in [25], cf. `mbsolve-tool/src/mbsolve-tool.cpp` in the mbsolve repository [119].

```
/* set up quantum mechanical descriptions */
auto qm_gain = std::make_shared<mbsolve::qm_desc_2lvl>(
  5e21,
  2 * mbsolve::PI * 3.4e12,
```

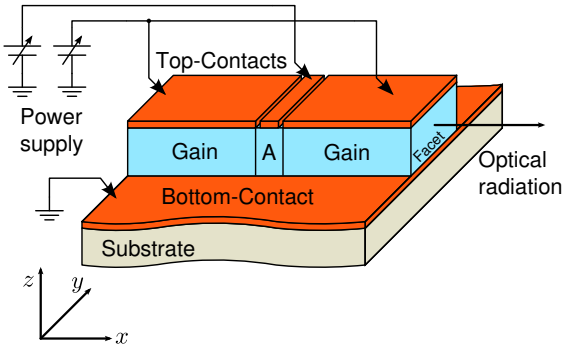


Figure 8.1 Schematic of the test case setup. The absorber region (A) is embedded between two gain regions. The region types are distinguished by the material and/or the bias voltage. Reprinted with permission from M. Riesch *et al.*, "Dynamic Simulations of Quantum Cascade Lasers Beyond the Rotating Wave Approximation" [103]. © 2018 IEEE.

```
  2e-9,
  1.0 / 10e-12,
  1.0 / 200e-15,
  1.0);

auto qm_absorber = std::make_shared<mbsolve::qm_desc_2lvl>(
  1e21,
  2 * mbsolve::PI * 3.4e12,
  6e-9,
  1.0 / 3e-12,
  1.0 / 160e-15);

/* materials */
auto mat_absorber = std::make_shared<mbsolve::material>(
  "Absorber", qm_absorber, 12.96, 1.0, 500);
auto mat_gain = std::make_shared<mbsolve::material>(
  "Gain", qm_gain, 12.96, 1.0, 500);
mbsolve::material::add_to_library(mat_absorber);
mbsolve::material::add_to_library(mat_gain);

/* set up device */
dev = std::make_shared<mbsolve::device>("tzenov-cpml");
dev->add_region(std::make_shared<mbsolve::region>(
  "Gain R", mat_gain, 0, 0.5e-3));
dev->add_region(std::make_shared<mbsolve::region>(
  "Absorber", mat_absorber, 0.5e-3, 0.625e-3));
dev->add_region(std::make_shared<mbsolve::region>(
  "Gain L", mat_gain, 0.625e-3, 1.125e-3));
```

The resulting full-wave simulation of 310 round trips takes approximately 20 minutes on an Intel i7-4790 desktop processor using four threads. The simulation results in Fig. 8.2a) show that two pulses per round trip are formed, each of them featuring an intensity FWHM duration of approximately 850 fs. This is in excellent agreement with the pulse envelopes from the RWA simulation code [25]. As already pointed out, the latter simulation yields only the envelope of the electric field. The envelope is subsequently multiplied with the carrier wave, which corresponds to the transition frequency ω_{21}. Then, also the spectrum of the radiation can be determined, which is depicted in Fig. 8.2b). We note that a frequency shift of 115 GHz is clearly visible between the two resulting spectra. This discrepancy can be attributed to different effects that are known to be omitted by the RWA, e.g., the Bloch-Siegert shift [177]. This demonstrates that certain effects that may be crucial in the analysis of the QCL dynamics are omitted when using the RWA, and that

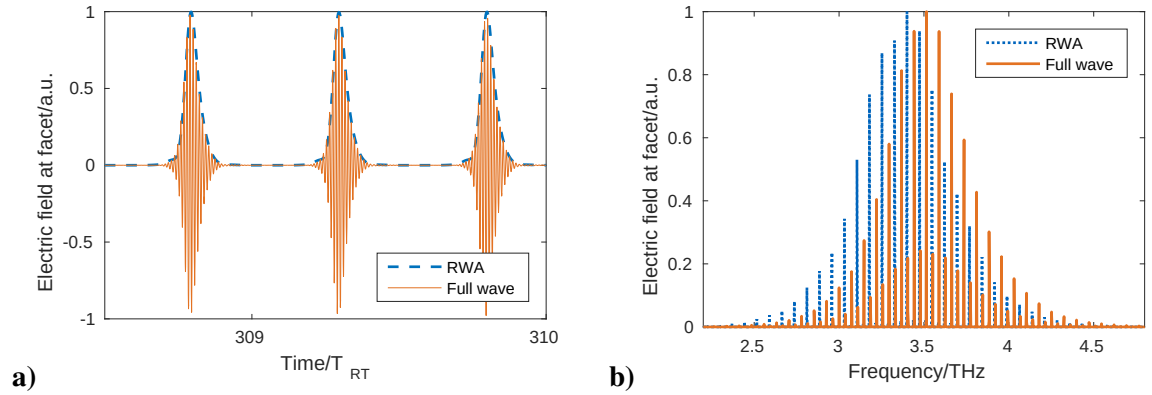


Figure 8.2 Comparison of the simulation results of the CPML setup with and without invoking the RWA. **a)** The electric field recorded at the right facet of the CPML setup. **b)** The amplitude spectrum of the electric field. Reprinted with permission from M. Riesch *et al.*, "Efficient Simulation of the Quantum Cascade Laser Dynamics beyond the Rotating Wave Approximation" [104].

the self-consistent full-wave approach may be required to avoid a priori assumptions about the resulting central frequency.

8.2 Modeling harmonic mode-locking in quantum cascade lasers

The time traces in Fig. 8.2a) exhibit the key feature of harmonic mode-locking, namely the emission of multiple pulses per round trip. In frequency domain, this corresponds to modes that are separated by a multiple of the round trip frequency f_{rt}, as can be seen in Fig. 8.2b). Harmonic mode-locking is routinely used in the visible and near-infrared regime for applications that require low noise [20]. Additionally, self-starting harmonic mode-locking has been observed in mid-infrared QCLs [178], and there are indications that terahertz QCLs could exhibit similar behavior [179]. In a recent study [20], time traces of the electromagnetic field in a passive harmonic mode-locking setup in the terahertz regime have been presented for the first time. While the time traces do not contain clearly separated pulses, their periodicity and quasi-pulsed nature strongly indicate that passive harmonic mode-locking in the terahertz regime is possible after all. In the frequency domain, the experimental results exhibit two significant features: the modes are separated by $2f_{\mathrm{rt}}$ and are grouped into two lobes, which are separated by $15f_{\mathrm{rt}}$. Correspondingly, there are 15 quasi-pulses in time domain per round trip, which can be grouped into two blocks. In the following, we aim to investigate the reason of these features.

To this end, we have to be aware that a radio frequency beat note signal can be measured in frequency comb setups (e.g., in [120], [180], [181]). This can be qualitatively explained by assuming that the polarization

$$P_z(E_z) = \varepsilon_0 \left[\chi^{(1)} E_z + \chi^{(2)} E_z^2 + \chi^{(3)} E_z^3 + \dots \right] \tag{8.1}$$

in Eqs. (3.67) depends on the electric field E_z in nonlinear fashion and can be expanded as power series. The electric field

$$E_z = \sum_k a_k \cos\left(\omega_{\mathrm{c}} t + k\omega_{\mathrm{rt}} t + \phi_k\right) \tag{8.2}$$

of a frequency comb, where we use a real-valued, discretized version of Eq. (1.3) in time domain, leads then to the term

$$E_z^2 = \sum_{jk} \frac{1}{2} a_j a_k \left\{ \cos\left[(j-k)\,\omega_{\mathrm{rt}} t + \phi_j - \phi_k\right] + \cos\left[2\omega_{\mathrm{c}} t + (j+k)\,\omega_{\mathrm{rt}} t + \phi_j + \phi_k\right] \right\}. \tag{8.3}$$

As we are only interested in the radio frequency component, we can ignore the terms with $\cos[2\omega_c t + \dots]$ and write

$$P_z^{(2)} = \varepsilon_0 \chi^{(2)} \left[\sum_{jk,k=j} \frac{1}{2} a_j^2 + \sum_{jk,k=j\pm1} \frac{1}{2} a_j a_k \cos\left(\omega_{\mathrm{rt}} t \pm \phi_j \mp \phi_k\right) + \sum_{jk,k=j\pm2} \frac{1}{2} a_j a_k \cos\left(2\omega_{\mathrm{rt}} t \pm \phi_j \mp \phi_k\right) + \dots \right] \tag{8.4}$$

for the second-order component of the polarization. This component enters Eqs. (3.67) as source term

$$\partial_t P_z^{(2)} = \sum_{l>0} b_l \cos\left(l\omega_{rt} t + \varphi_l\right), \tag{8.5}$$

where the amplitudes a_j and phase values ϕ_j have been aggregated to b_l and φ_l, respectively. As a consequence of its nonlinearity, the active region of QCLs may act as source of an additional electromagnetic signal. This signal oscillates with multiples of the round trip frequency, which is typically in the radio frequency domain, and can be measured with appropriate network analyzers.

The situation is quite similar in active mode-locking scenarios. Here, an external electromagnetic signal is injected into the QCL to achieve fundamental [180], [181] (signal oscillates with round trip frequency) or harmonic [20] (signal oscillates with the second harmonic of the round trip frequency) active mode-locking. Therefore, it is widely believed that the intrinsically generated radio frequency signal could have a similar impact on the QCL dynamics. In order to account for the radio frequency (or modulation) signal, the model has to be extended by a description of its source (intrinsic or external) as well as of its impact on the active region. This idea is not exactly new and first notions of this extension can be traced back to 2004 [182]. Additionally, the radio frequency properties of the resonator wave guide need to be considered (cf. [183], [184]), as they govern the propagation of the modulation signal in the QCL. Although there have been initial attempts to implement such an extension [185], [186], unfortunately no readily usable model exists at the moment.

Therefore, we choose a different path and consider the qualitative impact of an intrinsic modulation signal that oscillates with the second harmonic of the round trip frequency. Thereby we aim to confirm that such a signal could cause the features in the emitted terahertz radiation described above. This in turn would suggest the presence of such a second harmonic modulation signal, which could not be verified in [20] due to the lack of suitable measurement equipment. It has been reported that a modulation signal creates a population inversion grating [187]. As the population inversion is directly related to the gain and the QCLs are typically operated close to the lasing threshold, the modulation signal may cause regions that amplify or absorb the terahertz radiation, respectively. As the grating in [187] is time-dependent, there is no direct relation with the CPML setup in the previous section. However, using a mapping technique applied in [20] and described in [188] we can establish such a relation.

To this end, we consider a certain mode

$$E_m(x,t) = A_m \left\{ \sqrt{R} \exp\left[\frac{gx}{2} + \mathrm{i}\left(k_m x - \omega_m t\right)\right] + \exp\left[\frac{-gx}{2} + \mathrm{i}\left(k_m x + \omega_m t\right)\right] \right\} + \text{c.c.}, \tag{8.6}$$

of the modulation signal in the cavity (cf. Eq. (S5) in the Supplementary of [187]), where A_m is the mode amplitude, R represents the reflection at the facet, g denotes the net gain in the cavity, and k_m and $\omega_m = 2\pi m f_{\mathrm{rt}}$ are wave number and angular frequency of the mode, respectively. For simplicity, we assume a steady-state solution with perfectly reflecting mirrors (i.e., $R = 1$) and a perfectly balanced gain-loss ratio (i.e., the net gain $g = 0$). Then, the mode assumes a standing wave

$$E_m = B_m \cos\left(k_m x\right) \cos\left(\omega_m t\right), \tag{8.7}$$

where the amplitude $B_m = 4A_m$. Due to the periodicity of the cosine function we can replace the spatial coordinate $x \in [0, L_{\mathrm{cav}}]$ with a new coordinate $\xi = \xi_0 \pm ct/n \in (-\infty, \infty)$ that represents the position of a

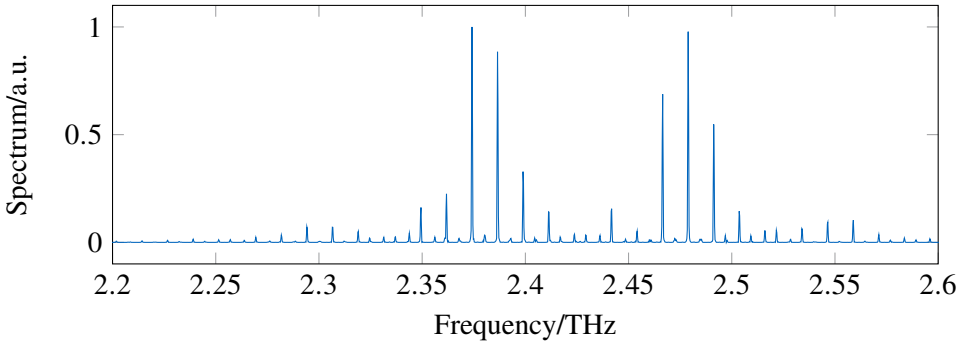


Figure 8.3 Simulated spectrum of the emitted radiation of the absorber-gain-absorber-gain-absorber setup. Reprinted from F. Wang *et al.*, "Ultrafast response of harmonic modelocked THz lasers" [20] (CC BY 4.0).

pulse moving with the velocity c/n. We plug the new coordinate into Eq. (8.7) and can thereby determine the modulation signal

$$\begin{aligned} E_m(\xi, t) &= B_m \cos\left(k_m \xi_0 \pm k_m c t/n\right) \cos\left(\omega_m t\right) \\ &= \frac{1}{2} B_m \left[\cos\left(k_m \xi_0 \pm k_m c t/n - \omega_m t\right) + \cos\left(k_m \xi_0 \pm k_m c t/n + \omega_m t\right)\right] \end{aligned} \tag{8.8}$$

that the moving pulse experiences. By further assuming that $k_m c/n = \omega_m$, i.e., the velocity values of the pulse and the modulation signal coincide, we can simplify the result to the expression

$$\begin{aligned} E_m(\xi, t) &= \frac{1}{2} B_m \left[\cos\left(k_m \xi_0\right) + \cos\left(k_m \xi_0 \pm 2 k_m c t/n\right)\right] \\ &= \frac{1}{2} B_m \left[\cos\left(k_m \xi_0\right) + \cos\left(-k_m \xi_0 + 2 k_m \xi\right)\right], \end{aligned} \tag{8.9}$$

which now only depends on the new spatial coordinate ξ. Again we exploit the periodicity of the cosine function and establish the fact that the pulse experiences the same modulation signal if it travels back and forth in a resonator cavity with the length L_{cav} and the corresponding time-independent modulation grating

$$E_m(x) = \frac{1}{2} B_m \left[\cos\left(k_m \xi_0\right) + \cos\left(-k_m \xi_0 + 2 k_m x\right)\right]. \tag{8.10}$$

We note that the direction in which the pulse travels is not relevant due to the symmetry of the cosine function. The initial pulse position ξ_0 can be considered as measure for the relation between the terahertz radiation and the modulation signal, which was found to be crucial in active mode-locking scenarios [180]. This is reflected by Eq. (8.10) as the modulation features an offset value $\cos(k_m \xi_0)$.

As already stated above, we assume that the QCL is operated slightly above its lasing threshold bias φ_{th}, e.g., at $\varphi_{\mathrm{th}} + B_m/2$. Then, the modulation grating (which is added to the bias) might drive the laser below threshold for $E_m < -B_m/2$. Therefore, the modulation grating defines whether a certain portion of the active region acts as gain or absorber medium. According to this reasoning, the complete active region acts as gain medium for $k_m \xi_0 = 0$. On the other hand, for $k_m \xi_0 = \pi$ the modulation grating

$$E_m(x) = -\frac{1}{2} B_m \left[1 + \cos\left(2 k_m x\right)\right] \tag{8.11}$$

clearly yields absorber and gain regions, similar to the CPML setup. In case of the device in [20], the grating leads to the sequence 0.75 mm absorber, 1.5 mm gain, 1.5 mm absorber, 1.5 mm gain, and 0.75 mm absorber, totaling up to a device length of 6 mm. The gain and absorber materials were implemented by selecting suitable parameters from a range of reasonable values, as was done e.g., in [23]. The resulting device can be readily simulated using mbsolve with a publicly available Python script [189].

The results of the simulation (depicted in Fig. 8.3) show that two features of the experimental results could be reproduced. The modes are separated by $2 f_{\mathrm{rt}}$ and can be grouped into two lobes, which are separated

by $15 f_{\mathrm{rt}}$. While the overall agreement with the experimental results leaves room for improvement, the corresponding features in time domain are also visible (cf. [20]). This strongly suggests that an intrinsic modulation signal oscillating with $2 f_{\mathrm{rt}}$ can be responsible for those effects. We need to keep in mind, however, that this is only a qualitative analysis and we are still far from a self-consistent, quantitative model.

8.3 Dynamics of a quantum cascade laser with a graphene reflector

This simulation setup bases on the work in [173]. In this recent study a QCL was coupled with a graphene-coated mirror, thereby forming an external resonator cavity. Most notably, it was shown that the 65 nm thin graphene layer caused a significant improvement in the linewidth of the beat note signal. This was demonstrated with the help of a similar setup that used an uncoated mirror. Here, the beat note linewidth increased by five orders of magnitude (cf. Fig. 6 in [173]). In addition to the latter experiment, simulations were carried out to investigate whether the graphene layer could indeed have such impact on the beat note linewidth.

To this end, the active region of the QCL is modeled as two-level gain medium with empirically chosen parameters, selected from a range of reasonable values, to match the measured experimental laser spectrum. The graphene material is assumed as saturable absorber, where the power loss coefficient

$$\alpha(I) = \alpha_{\mathrm{NS}} + \frac{\alpha_{\mathrm{S}}}{1 + (I/I_{\mathrm{S}})} \tag{8.12}$$

depends on the intensity I [190]. Here, I_{S} is the saturation intensity, and α_{S} and α_{NS} are the saturable and non-saturable loss coefficient, respectively. The non-saturable loss coefficient can be converted to a linear amplitude loss α_0 (as used in mbsolve), and the saturable part can be generally described as a two-level system [32], [83], [191] using the relations

$$\alpha_{\mathrm{S}} = \Gamma \omega_{21} n_{3\mathrm{D}} \frac{|d_{21}|^2 T_2}{n_{\mathrm{eff}}} \frac{1}{\varepsilon_0 c \hbar} \tag{8.13}$$

and

$$I_{\mathrm{S}} = \hbar^2 c \varepsilon_0 \frac{n_{\mathrm{eff}}}{|z_{21}|^2 T_2} \frac{1}{2T_1}. \tag{8.14}$$

After invoking the adiabatic elimination and assuming instantaneous absorption (i.e., $T_1, T_2 \approx 0$), as well as using the values from the QCL model for the background effective refractive index, the overlap factor, and the transition frequency, the values for the dipole length z_{21} and the charge carrier density $n_{3\mathrm{D}}$ can be readily determined.

The device setup in the simulation neglects the air gap, as it is significantly smaller than the central wavelength of the emitted radiation and its effect on the operation is limited to loss, independent of frequency and intensity. The air gap, and similarly the outcoupling of radiation, is considered in the form of an increased loss in the QCL resonator wave guide. The ultra-thin graphene layer constitutes a major challenge for the finite-difference time-domain method, as the spatial discretization size must be reduced significantly. In order to maintain computational efficiency, a further approximation is used. It is assumed that the main effect of the graphene is the reflection at the interface, and that the radiation that enters the graphene layer is rapidly attenuated. Then, we can artificially extend the dimension of the absorber, thus allowing a larger spatial discretization size. In a second simulation run, the graphene layer can be replaced by a perfectly reflecting mirror to extract its impact on the device operation.

As next step, the setup can be simulated with mbsolve and the resulting recorded time traces of the emitted radiation can be transformed to frequency domain for comparison with the measured spectrum. Two methods are available to determine the beat note signal. One approach follows Eq. (8.1) and relates the beat note signal to the square of the electric field, which is subsequently transformed to frequency domain [23]. Alternatively, the auto-correlation ansatz used e.g., in [173], [188] can be employed. Although limited in spectral resolution, the simulations with and without the graphene layer clearly confirm that the graphene layer is responsible for the significant reduction [173].

8.4 Summary

We have seen in this chapter that the simulation framework presented in the thesis at hand is already able to aid theoretical and experimental studies. The flexibility of the framework already facilitates the modeling and simulation of various experimental setups, and allows the accurate reproduction of theoretical results without invoking the RWA. Here, a particularly interesting feature is the ability to model multi-section setups. A first approach to integrate different QCL active regions laterally has been published [192], opening up a variety of possible applications. Also, the interaction with materials outside of the QCL (cf. the graphene-coated mirror in Section 8.3) can be simulated appropriately.

Nevertheless, several challenges remain to be mastered. While qualitative models for the impact of the beat note signal in Section 8.2 and the generation of the beat note signal in Section 8.3 could confirm the respective working hypothesis, we note that quantitative and self-consistent models are required to predict the dynamical behavior of a certain setup. Here, the simulation framework at hand provides a solid basis to build upon.

9 Conclusion and outlook

Lately it occurs to me what a long strange trip it's been.
— *Grateful Dead, Truckin'*

Since the application range of quantum cascade lasers does not depend on the performance (with respect to output power and/or operating temperature) alone but also on the availability of different operating modes, the computer-aided investigation of the dynamical behavior of QCLs has been in the focus of the thesis at hand. As we have seen, this investigation requires the discussion of several topics spanning different fields: physics, (numerical) mathematics, and computer science. In the following, we conclude this thesis by iterating over these topics, where a short summary as well as an outlook on possible future tasks is given.

Every computational investigation in physics needs to base on a solid theoretical foundation. In the work at hand, this foundation has been provided in Chapter 2 in the form of a compact introduction to classical electrodynamics as well as quantum mechanics and its application to semiconductors. Basing on this introduction, the underlying model has been derived in Chapter 3. In this rigorous derivation care has been taken to consider and describe all assumptions and approximations. Although the resulting model is commonly used in related work, such a thorough description has not yet been published to the best of the author's knowledge. As pointed out in Chapter 8, which has presented the application of the model to different experimental setups, and in related literature, good agreement with experimental results can be achieved using the model at hand.

However, certain features in experimental results cannot be explained with the current model. In many experimental setups a radio frequency probe is attached to the contacts of a QCL and the electromagnetic signal is measured with a network analyzer. As discussed in Chapter 8, a beat note signal can be frequently measured and indicates frequency comb operation. The generation of this beat note signal is not accounted for in state-of-the-art models. It has been indicated that the nonlinear polarization current density $\partial_t P_z$ that stems from the quantum mechanical systems is the source of this signal, in which case the model would already include the beat note generation intrinsically. However, due to the different frequency regimes the beat note signal is bound to have different propagation properties as the optical radiation field. For example, the difference in the refractive index as well as its consequences have been discussed in [184], [188]. As a consequence, an additional set of Maxwell's equations for the beat note signal with different electromagnetic properties has been suggested in [182], [185], [186]. While different forms are possible (e.g., the telegrapher's equations), it is probably most straight-forward to write

$$\partial_x E_{\mathrm{RF},z} = \mu_0 \partial_t H_{\mathrm{RF},y}, \tag{9.1a}$$

$$\partial_x H_{\mathrm{RF},y} = -\varepsilon_0 \varepsilon_{\mathrm{RF,eff}} \partial_t E_{\mathrm{RF},z} - \sigma_{\mathrm{RF,eff}} E_{\mathrm{RF},z} - \Gamma_{\mathrm{RF}} \partial_t P_{\mathrm{RF},z}, \tag{9.1b}$$

for the radio-frequency portion of the electromagnetic field, in analogy to Eq. (3.67). Here, $\partial_t P_{\mathrm{RF},z}$ represents the radio-frequency contribution to the polarization current density, which needs to be extracted from the total current density. While the implementation of a second set of Maxwell's equations is trivial, the exact extraction model as well as the calculation of the radio-frequency electromagnetic properties ($\varepsilon_{\mathrm{RF,eff}}$, $\sigma_{\mathrm{RF,eff}}$, and Γ_{RF}) need further discussion. As a subsequent step, the impact of the radio-frequency signal on the QCL operation should be modeled. Thereby, the interaction between the radio-frequency signal and the QCL operation would be modeled completely. Apart from QCLs the resulting traveling wave model could be applicable to other quantum devices, such as Josephson junctions [193]. In contrast to the optical radiation field the radio-frequency signal cannot be treated as perturbation (at least in some cases), but it may help to assume that the signal is slowly varying when compared to the transition frequencies between the energy levels. In the end, the quantum mechanical operators $\hat{H}_0$ and $\hat{\mu}_z$ as well as the dissipation superoperator

$\mathcal{D}$ may depend on the radio-frequency electric field $E_{\mathrm{RF},z}$. Finally, soft sources could be added in the numerical treatment of Eqs. (9.1) to model external radio-frequency sources, which are commonly used for active mode-locking of QCLs. Such an extended model for the QCL dynamics would be helpful to simulate active mode-locking setups and to investigate the self-starting effects in harmonic mode-locking setups, which are discussed in e.g. [20], [194].

After the mathematical model is defined, the second step of a computational investigation is typically to find suitable numerical methods for the involved differential equations. Since for the static prerequisite models (cf. Fig. 3.1) numerical methods and tools exist already, the discussion in Chapter 6 has focused on the treatment of the Maxwell-Bloch equations including methods for Maxwell's equations, various numerical techniques for the Lindblad equation, and the coupling between them. Apart from suitable and ready-to-use methods for the Maxwell-Bloch equations promising alternatives have been discussed. Here, the algorithm by Al-Mohy and Higham [153] for solving the Lindblad equation deserves explicit mention, as it features the largest accuracy and reasonable performance. Seeing that the model might be extended in near future to account for the radio-frequency signal, which leads to more time-dependent quantities in the Lindblad equation, the approach shall be even more helpful. Since it does not depend on precalculated quantities its performance will not decrease in this case. The absolute performance may be increased by further mathematical optimizations (such as using improved parameter calculations [195]), which should be in the focus of future work. As far as Maxwell's equations are concerned, the finite-difference time-domain (FDTD) method has proven to be well-suited for the use in QCL simulations. However, the inclusion of dispersion is significantly more complex in comparison to the methods that invoke the rotating-wave approximation. Since dispersion effects play an essential role in the QCL dynamics, the FDTD should be extended to allow for frequency dependent electromagnetic properties. Fortunately, suitable numerical approaches (e.g., in [196]) are already available and there is even a draft implementation in the scope of the mbsolve project [197].

Following the numerical treatment, the implementation of a (ideally publicly available) solver tool constitutes the next step of a computational investigation. In order to provide a sustainable solution, certain considerations with respect to the software development process are required. Chapter 5 discusses in detail those considerations and presents a skeleton project that can be used as template for new or as guideline for existing scientific software projects. In the case of mbsolve, the open-source solver for the Maxwell-Bloch equations presented in Chapter 4, the skeleton project was used as guideline to provide a software package that is easy to install and use. The flexible interface (in C++ or Python) does not only allow the convenient usage of the simulation tool, but could also enable the inclusion in automated optimization tools for quantum cascade lasers, such as those presented in [198], [199]. The required computational performance is achieved by distributing the workload among many processors using the OpenMP standard. This approach is discussed in Chapter 7. While for the current simulation tasks the OpenMP approach for shared memory systems is sufficient, future simulations may require the usage of larger high-performance computing systems. Therefore, the support for distributed memory systems should be in the focus of future work. Also, many additional features can be envisaged, such as support for periodic spatial boundary conditions to model ring cavities.

Although there are some tasks left for future work, we note that the main goal of this thesis has been achieved: the envisaged flexible software framework for the computer-aided simulation of the dynamical behavior of quantum cascade lasers has been implemented and is publicly available. As a side effect, the solver tool can be used for the investigation of other quantum mechanical effects (outside the scope of QCL simulations), such as Rabi flopping [200], and the modeling of other nano-optoelectronic structures and devices (e.g., based on quantum dots). While the application of the solver tool to actual problems constitutes a relatively small part of the thesis at hand, the theoretical description, the numerical treatment, and the implementation tasks carried out shall provide a solid basis for upcoming simulation work.

Bibliography

[1] R. F. Kazarinov and R. A. Suris, "Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice", *Sov. Phys. Semicond.*, vol. 5, no. 4, pp. 797–800, Oct. 1971.

[2] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser", *Science*, vol. 264, no. 5158, pp. 553–556, Apr. 1994. DOI: 10.1126/science.264.5158.553.

[3] M. S. Vitiello, G. Scalari, B. Williams, and P. De Natale, "Quantum cascade lasers: 20 years of challenges", *Opt. Express*, vol. 23, no. 4, pp. 5167–5182, Feb. 2015. DOI: 10.1364/OE.23.005167.

[4] A. Müller and J. Faist, "The quantum cascade laser: Ready for take-off", *Nat. Photon.*, vol. 4, no. 5, p. 291, May 2010. DOI: 10.1038/nphoton.2010.110.

[5] C. Sirtori, "Terahertz race heats up", *Nat. Photonics*, vol. 15, no. 1, pp. 1–2, Jan. 2021. DOI: 10.1038/s41566-020-00740-4.

[6] S. Dhillon, M. Vitiello, E. Linfield, A. Davies, M. C. Hoffmann, J. Booske, C. Paoloni, M. Gensch, P. Weightman, G. Williams, *et al.*, "The 2017 terahertz science and technology roadmap", *J. Phys. D: Appl. Phys.*, vol. 50, no. 4, p. 043 001, Jan. 2017. DOI: 10.1088/1361-6463/50/4/043001.

[7] 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*, vol. 417, no. 6885, pp. 156–159, May 2002. DOI: 10.1038/417156a.

[8] J.-F. Lampin, G. Mouret, S. Dhillon, and J. Mangeney, "THz spectroscopy for fundamental science and applications", *Photoniques*, no. 101, pp. 33–38, Apr. 2020. DOI: 10.1051/photon/202010133.

[9] C. Jirauschek and T. Kubis, "Modeling techniques for quantum cascade lasers", *Appl. Phys. Rev.*, vol. 1, p. 011 307, Feb. 2014. DOI: 10.1063/1.4863665.

[10] A. E. Siegman, *Lasers*. Mill Valley, California: University Science Books, 1986.

[11] T. W. Hänsch, "Nobel lecture: Passion for precision", *Rev. Mod. Phys.*, vol. 78, no. 4, p. 1297, Nov. 2006. DOI: 10.1103/RevModPhys.78.1297.

[12] D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 7th ed. John Wiley & Sons, New York, 2005.

[13] T. Fortier and E. Baumann, "20 years of developments in optical frequency comb technology and applications", *Commun. Phys.*, vol. 2, no. 1, pp. 1–16, Dec. 2019. DOI: 10.1038/s42005-019-0249-y.

[14] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed. McGraw-Hill, Boston, 2000.

[15] H. A. Haus, "Theory of mode locking with a fast saturable absorber", *J. Appl. Phys.*, vol. 46, no. 7, pp. 3049–3058, Jul. 1975. DOI: 10.1063/1.321997.

[16] H. A. Haus, "Mode-locking of lasers", *IEEE J. Sel. Top. Quant.*, vol. 6, no. 6, pp. 1173–1185, Nov. 2000. DOI: 10.1109/2944.902165.

[17] M. Riesch, *qclsip: The Quantum Cascade Laser Stock Image Project*, Jan. 2019. DOI: 10.5281/zenodo.2641239.

[18] 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 ~200 K with optimized oscillator strength and improved injection tunneling", *Opt. Express*, vol. 20, no. 4, pp. 3866–3876, Feb. 2012. DOI: 10.1364/OE.20.003866.

[19] G. Villares, A. Hugi, S. Blaser, and J. Faist, "Dual-comb spectroscopy based on quantum-cascade-laser frequency combs", *Nat. Commun.*, vol. 5, p. 5192, Oct. 2014. DOI: 10.1038/ncomms6192.

[20] F. Wang, V. Pistore, M. Riesch, H. Nong, P.-B. Vigneron, R. Colombelli, O. Parillaud, J. Mangeney, J. Tignon, C. Jirauschek, and S. S. Dhillon, "Ultrafast response of active and self-starting harmonic modelocked THz laser", *Light Sci. Appl.*, vol. 9, no. 1, p. 51, Apr. 2020. DOI: 10.1038/s41377-020-0288-x.

[21] C. Y. Wang, L. Diehl, A. Gordon, C. Jirauschek, F. Kärtner, A. Belyanin, D. Bour, S. Corzine, G. Höfler, M. Troccoli, J. Faist, and F. Capasso, "Coherent instabilities in a semiconductor laser with fast gain recovery", *Phys. Rev. A*, vol. 75, no. 3, p. 031 802, Mar. 2007. DOI: 10.1103/PhysRevA.75.031802.

[22] V.-M. Gkortsas, C. Y. Wang, L. Kuznetsova, L. Diehl, A. Gordon, C. Jirauschek, M. Belkin, A. Belyanin, F. Capasso, and F. Kärtner, "Dynamics of actively mode-locked quantum cascade lasers", *Opt. Express*, vol. 18, no. 13, pp. 13 616–13 630, Jun. 2010. DOI: 10.1364/OE.18.013616.

[23] P. Tzenov, D. Burghoff, Q. Hu, and C. Jirauschek, "Time domain modeling of terahertz quantum cascade lasers for frequency comb generation", *Opt. Express*, vol. 24, no. 20, pp. 23 232–23 247, Sep. 2016. DOI: 10.1364/OE.24.023232.

[24] J. R. Freeman, J. Maysonnave, S. Khanna, E. H. Linfield, A. G. Davies, S. Dhillon, and J. Tignon, "Laser-seeding dynamics with few-cycle pulses: Maxwell-Bloch finite-difference time-domain simulations of terahertz quantum cascade lasers", *Phys. Rev. A*, vol. 87, no. 6, p. 063 817, Jun. 2013. DOI: 10.1103/PhysRevA.87.063817.

[25] P. Tzenov, I. Babushkin, R. Arkhipov, M. Arkhipov, N. N. Rosanov, U. Morgner, and C. Jirauschek, "Passive and hybrid mode locking in multi-section terahertz quantum cascade lasers", *New J. Phys.*, vol. 20, no. 5, p. 053 055, May 2018. DOI: 10.1088/1367-2630/aac12a.

[26] D. J. Griffiths, *Introduction to Electrodynamics*, 5th ed. Cambridge University Press, Cambridge, 2017.

[27] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. John Wiley & Sons, New York, 1999.

[28] J.-M. Jin, *Theory and Computation of Electromagnetic Fields*, 2nd ed. John Wiley & Sons, New York, 2015.

[29] C. L. Tang, *Fundamentals of Quantum Mechanics: For Solid State Electronics and Optics*. Cambridge University Press, Cambridge, 2005.

[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. Cambridge University Press, Cambridge, 2010.

[31] P. Meystre and M. Sargent, *Elements of Quantum Optics*, 4th ed. Springer Sciene & Business Media, New York, 2007.

[32] C. Jirauschek, M. Riesch, and P. Tzenov, "Optoelectronic device simulations based on macroscopic Maxwell–Bloch equations", *Adv. Theor. Simul.*, vol. 2, no. 8, p. 1 900 018, Aug. 2019. DOI: 10.1002/adts.201900018.

[33] M. Göppert-Mayer, "Über Elementarakte mit zwei Quantensprüngen", *Ann. Phys.*, vol. 401, no. 3, pp. 273–294, 1931. DOI: 10.1002/andp.19314010303.

[34] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, 2002.

[35] M. M. Wilde, *Quantum Information Theory*, 2nd ed. Cambridge University Press, Cambridge, 2017.

[36] M.-D. Choi, "Completely positive linear maps on complex matrices", *Linear Algebra Its Appl.*, vol. 10, no. 3, pp. 285–290, Jun. 1975.

[37] K. Kraus, "General state changes in quantum theory", *Ann. Phys.*, vol. 64, no. 2, pp. 311–335, Jun. 1971.

[38] G. Lindblad, "On the generators of quantum dynamical semigroups", *Commun. Math. Phys.*, vol. 48, no. 2, pp. 119–130, Jun. 1976. DOI: 10.1007/BF01608499.

[39] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, "Completely positive dynamical semigroups of *N*-level systems", *J. Math. Phys.*, vol. 17, no. 5, pp. 821–825, May 1976. DOI: 10.1063/1.522979.

[40] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Holt, Rinehart and Winston, New York, 1976.

[41] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors: Physics and Materials Properties*, 4th ed. Springer-Verlag, Berlin Heidelberg, 2010.

[42] J. H. Davies, *The Physics of Low-dimensional Semiconductors: An Introduction*. Cambridge University Press, Cambridge, 1998.

[43] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*. Les éditions de physique, Les Ulis, 1988.

[44] G. Floquet, "Sur les équations différentielles linéaires à coefficients périodiques", *Ann. de l'Ecole Norm. Sup.*, vol. 12, pp. 47–88, 1883. DOI: 10.24033/asens.220.

[45] F. Bloch, "Über die Quantenmechanik der Elektronen in Kristallgittern", *Z. Physik*, vol. 52, pp. 555–600, Jul. 1929. DOI: 10.1007/BF01339455.

[46] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Springer, Wien, 1989.

[47] P. Borowik, J.-L. Thobel, and L. Adamowicz, "Monte Carlo modeling applied to studies of quantum cascade lasers", *Opt. Quant. Electron.*, vol. 49, no. 3, p. 96, Feb. 2017. DOI: 10.1007/s11082-017-0931-9.

[48] R. C. Iotti, E. Ciancio, and F. Rossi, "Quantum transport theory for semiconductor nanostructures: A density-matrix formulation", *Phys. Rev. B*, vol. 72, no. 12, p. 125 347, Sep. 2005. DOI: 10.1103/PhysRevB.72.125347.

[49] G. Kiršanskas, M. Franckié, and A. Wacker, "Phenomenological position and energy resolving Lindblad approach to quantum kinetics", *Phys. Rev. B*, vol. 97, no. 3, p. 035 432, Jan. 2018. DOI: 10.1103/PhysRevB.97.035432.

[50] S. Soleimanikahnoj, O. Jonasson, F. Karimi, and I. Knezevic, "Numerically efficient density-matrix technique for modeling electronic transport in mid-infrared quantum cascade lasers", *J. Comput. Electron.*, Jan. 2021. DOI: 10.1007/s10825-020-01627-x.

[51] A. Wacker, "Semiconductor superlattices: A model system for nonlinear transport", *Phys. Rep.*, vol. 357, no. 1, pp. 1–111, Jan. 2002. DOI: 10.1016/S0370-1573(01)00029-1.

[52] A. Wacker, M. Lindskog, and D. O. Winge, "Nonequilibrium Green's function model for simulation of quantum cascade laser devices under operating conditions", *IEEE J. Sel. Top. Quant.*, vol. 19, no. 5, pp. 1–11, Jan. 2013. DOI: 10.1109/JSTQE.2013.2239613.

[53] M. Franckié, L. Bosco, M. Beck, C. Bonzon, E. Mavrona, G. Scalari, A. Wacker, and J. Faist, "Two-well quantum cascade laser optimization by non-equilibrium Green's function modelling", *Appl. Phys. Lett.*, vol. 112, no. 2, p. 021 104, Jan. 2018. DOI: 10.1063/1.5004640.

[54] C. Jirauschek, "Density matrix Monte Carlo modeling of quantum cascade lasers", *J. Appl. Phys.*, vol. 122, no. 13, p. 133 105, Oct. 2017. DOI: 10.1063/1.5005618.

[55] C. G. Derntl, G. Scalari, D. Bachmann, M. Beck, J. Faist, K. Unterrainer, and J. Darmo, "Gain dynamics in a heterogeneous terahertz quantum cascade laser", *Appl. Phys. Lett.*, vol. 113, no. 18, p. 181 102, Oct. 2018. DOI: 10.1063/1.5049384.

[56] J. Faist, *Quantum Cascade Lasers*. Oxford University Press, Oxford, 2013.

[57] S. G. Schirmer and A. I. Solomon, "Constraints on relaxation rates for N-level quantum systems", *Phys. Rev. A*, vol. 70, p. 022 107, Aug. 2004. DOI: 10.1103/PhysRevA.70.022107.

[58] D. K. L. Oi and S. G. Schirmer, "Limits on the decay rate of quantum coherence and correlation", *Phys. Rev. A*, vol. 86, p. 012 121, Jul. 2012. DOI: 10.1103/PhysRevA.86.012121.

[59] C. Jirauschek and P. Tzenov, "Self-consistent simulations of quantum cascade laser structures for frequency comb generation", *Opt. Quant. Electron.*, vol. 49, no. 12, p. 414, Nov. 2017. DOI: 10.1007/s11082-017-1253-7.

[60] T. Ando, "Broadening of inter-subband transitions in image-potential-induced surface states outside liquid Helium", *J. Phys. Soc. Jpn.*, vol. 44, no. 3, pp. 765–773, Mar. 1978. DOI: 10.1143/JPSJ.44.765.

[61] T. Ando, "Line width of inter-subband absorption in inversion layers: Scattering from charged ions", *J. Phys. Soc. Jpn.*, vol. 54, no. 7, pp. 2671–2675, Jul. 1985. DOI: 10.1143/JPSJ.54.2671.

[62] T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, "Intersubband absorption linewidth in GaAs quantum wells due to scattering by interface roughness, phonons, alloy disorder, and impurities", *J. Appl. Phys.*, vol. 93, no. 3, pp. 1586–1597, Feb. 2003. DOI: 10.1063/1.1535733.

[63] E. Cassan, "On the reduction of direct tunneling leakage through ultrathin gate oxides by a one-dimensional Schrödinger–Poisson solver", *J. Appl. Phys.*, vol. 87, no. 11, pp. 7931–7939, May 2000. DOI: 10.1063/1.373477.

[64] H. Li, J. C. Cao, and H. C. Liu, "Effects of design parameters on the performance of terahertz quantum-cascade lasers", *Semicond. Sci. Technol.*, vol. 23, no. 12, p. 125 040, Nov. 2008. DOI: 10.1088/0268-1242/23/12/125040.

[65] E. Schrödinger, "Quantisierung als Eigenwertproblem", *Ann. Phys.*, vol. 385, no. 13, pp. 437–490, 1926. DOI: 10.1002/andp.19263851302.

[66] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. Academic Press, Boston, 2001.

[67] M. Riesch and C. Jirauschek, "Open-source simulation software for quantum cascade lasers", in *Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2019, 44th International Conference on*, Paris, France: IEEE, Piscataway, NJ, Sep. 2019. DOI: 10.1109/IRMMW-THz.2019.8874442.

[68] C. Jirauschek, "Accuracy of transfer matrix approaches for solving the effective mass Schrödinger equation", *IEEE J. Quantum Electron.*, vol. 45, no. 9, pp. 1059–1067, Jul. 2009. DOI: 10.1109/JQE.2009.2020998.

[69] C. Juang, K. J. Kuhn, and R. B. Darling, "Stark shift and field-induced tunneling in $Al_xGa_{1-x}As$/GaAs quantum-well structures", *Phys. Rev. B*, vol. 41, no. 17, pp. 12 047–12 053, Jun. 1990. DOI: 10.1103/PhysRevB.41.12047.

[70] H. Hebal, Z. Koziol, S. B. Lisesivdin, and R. Steed, "General-purpose open-source 1D self-consistent Schrödinger-Poisson solver: Aestimo 1D", *Comput. Mater. Sci.*, vol. 186, p. 110 015, Jan. 2021. DOI: https://doi.org/10.1016/j.commatsci.2020.110015.

[71] R. Steed, H. Hebal, and S. B. Lisesivdin, *Sblisesivdin/aestimo: Version v.1.2*, https://doi.org/10.5281/zenodo.1042657, Nov. 2017. DOI: 10.5281/zenodo.1042657.

[72] A. S. Sudbø, "Film mode matching: A versatile numerical method for vector mode field calculations in dielectric waveguides", *Pure Appl. Opt.*, vol. 2, no. 3, pp. 211–233, May 1993. DOI: 10.1088/0963-9659/2/3/007.

[73] M. Lohmeyer, "Wave-matching method for mode analysis of dielectric waveguides", *Opt. Quant. Electron.*, vol. 29, no. 9, pp. 907–922, Sep. 1997. DOI: 10.1023/A:1018581701193.

[74] M. Hammer, *Wave-matching method mode solver*, https://wmm.computational-photonics.eu/.

[75] C. R. Menyuk and M. A. Talukder, "Self-induced transparency modelocking of quantum cascade lasers", *Phys. Rev. Lett.*, vol. 102, no. 2, p. 023 903, Jan. 2009. DOI: 10.1103/PhysRevLett.102.023903.

[76] H. Choi, V.-M. Gkortsas, L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, F. Capasso, F. X. Kärtner, and T. B. Norris, "Ultrafast Rabi flopping and coherent pulse propagation in a quantum cascade laser", *Nat. Photonics*, vol. 4, no. 10, pp. 706–710, Aug. 2010. DOI: 10.1038/nphoton.2010.205.

[77] M. A. Talukder and C. R. Menyuk, "Quantum coherent saturable absorption for mid-infrared ultrashort pulses", *Opt. Express*, vol. 22, no. 13, pp. 15 608–15 617, Jun. 2014. DOI: 10.1364/OE.22.015608.

[78] Y. Wang and A. Belyanin, "Active mode-locking of mid-infrared quantum cascade lasers with short gain recovery time", *Opt. Express*, vol. 23, no. 4, pp. 4173–4185, Feb. 2015. DOI: 10.1364/OE.23.004173.

[79] N. N. Vuković, J. Radovanović, V. Milanović, and D. L. Boiko, "Low-threshold RNGH instabilities in quantum cascade lasers", *IEEE J. Sel. Top. Quant.*, vol. 23, no. 6, pp. 1–16, May 2017. DOI: 10.1109/JSTQE.2017.2699139.

[80] L. Columbo, S. Barbieri, C. Sirtori, and M. Brambilla, "Dynamics of a broad-band quantum cascade laser: From chaos to coherent dynamics and mode-locking", *Opt. Express*, vol. 26, no. 3, pp. 2829–2847, Jan. 2018. DOI: 10.1364/OE.26.002829.

[81] M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations", *Comput. Phys. Commun.*, vol. 268, p. 108 097, Nov. 2021. DOI: 10.1016/j.cpc.2021.108097.

[82] L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms*. Dover Publications, Mineola, NY, 1987.

[83] R. W. Boyd, *Nonlinear Optics*, 3rd ed. Academic Press, 2008.

[84] F. Bloch, "Nuclear induction", *Phys. Rev.*, vol. 70, pp. 460–474, 7-8 Oct. 1946. DOI: 10.1103/PhysRev.70.460.

[85] R. P. Feynman, F. L. Vernon, and R. W. Hellwarth, "Geometrical representation of the Schrödinger equation for solving maser problems", *J. Appl. Phys.*, vol. 28, no. 1, pp. 49–52, Sep. 1957. DOI: 10.1063/1.1722572.

[86] F. T. Arecchi and R. Bonifacio, "Theory of optical maser amplifiers", *IEEE J. Quantum Electron.*, vol. 1, no. 4, pp. 169–178, Jul. 1965. DOI: 10.1109/JQE.1965.1072212.

[87] I. D. Abella, N. A. Kurnit, and S. R. Hartmann, "Photon echoes", *Phys. Rev.*, vol. 141, pp. 391–406, 1 Jan. 1966. DOI: 10.1103/PhysRev.141.391.

[88] S. L. McCall and E. L. Hahn, "Self-induced transparency by pulsed coherent light", *Phys. Rev. Lett.*, vol. 18, pp. 908–911, 21 May 1967. DOI: 10.1103/PhysRevLett.18.908.

[89] S. L. McCall and E. L. Hahn, "Self-induced transparency", *Phys. Rev.*, vol. 183, pp. 457–485, 2 Jul. 1969. DOI: 10.1103/PhysRev.183.457.

[90] F. T. Hioe and J. H. Eberly, "N-level coherence vector and higher conservation laws in quantum optics and quantum mechanics", *Phys. Rev. Lett.*, vol. 47, no. 12, pp. 838–841, Sep. 1981. DOI: 10.1103/PhysRevLett.47.838.

[91] L. V. Hau, S. E. Harris, Z. Dutton, and C. H. Behroozi, "Light speed reduction to 17 metres per second in an ultracold atomic gas", *Nature*, vol. 397, no. 6720, pp. 594–598, 1999. DOI: 10.1038/17561.

[92] C. Liu, Z. Dutton, C. H. Behroozi, and L. V. Hau, "Observation of coherent optical information storage in an atomic medium using halted light pulses", *Nature*, vol. 409, no. 6819, pp. 490–493, 2001. DOI: 10.1038/35054017.

[93] R. W. Ziolkowski, J. M. Arnold, and D. M. Gogny, "Ultrafast pulse interactions with two-level atoms", *Phys. Rev. A*, vol. 52, no. 4, pp. 3082–3094, Oct. 1995. DOI: 10.1103/PhysRevA.52.3082.

[94] W. Cartar, J. Mørk, and S. Hughes, "Self-consistent Maxwell-Bloch model of quantum-dot photonic-crystal-cavity lasers", *Phys. Rev. A*, vol. 96, no. 2, p. 023 859, Aug. 2017. DOI: 10.1103/PhysRevA.96.023859.

[95] P. Gładysz, P. Wcisło, and K. Słowik, "Propagation of optically tunable coherent radiation in a medium of asymmetric molecules", *Sci. Rep.*, vol. 10, no. 1, p. 17 615, Oct. 2020. DOI: 10.1038/s41598-020-74569-w.

[96] G. Slavcheva, J. M. Arnold, I. Wallace, and R. W. Ziolkowski, "Coupled Maxwell-pseudospin equations for investigation of self-induced transparency effects in a degenerate three-level quantum system in two dimensions: Finite-difference time-domain study", *Phys. Rev. A*, vol. 66, no. 6, p. 63 418, Dec. 2002. DOI: 10.1103/PhysRevA.66.063418.

[97] M. Sukharev and A. Nitzan, "Numerical studies of the interaction of an atomic sample with the electromagnetic field in two dimensions", *Phys. Rev. A*, vol. 84, no. 4, p. 043 802, Oct. 2011. DOI: 10.1103/PhysRevA.84.043802.

[98] R. Marskar and U. Österberg, "Multilevel Maxwell-Bloch simulations in inhomogeneously broadened media", *Opt. Express*, vol. 19, no. 18, pp. 16 784–16 796, Aug. 2011. DOI: 10.1364/OE.19.016784.

[99] P. K. Nielsen, H. Thyrrestrup, J. Mørk, and B. Tromborg, "Numerical investigation of electromagnetically induced transparency in a quantum dot structure", *Opt. Express*, vol. 15, no. 10, pp. 6396–6408, May 2007. DOI: 10.1364/OE.15.006396.

[100] N. Majer, K. Lüdge, and E. Schöll, "Cascading enables ultrafast gain recovery dynamics of quantum dot semiconductor optical amplifiers", *Phys. Rev. B*, vol. 82, no. 23, p. 235 301, Dec. 2010. DOI: 10.1103/PhysRevB.82.235301.

[101] G. Slavcheva, M. Koleva, and A. Rastelli, "Ultrafast pulse phase shifts in a charged-quantum-dot–micropillar system", *Phys. Rev. B*, vol. 99, no. 11, p. 115 433, Mar. 2019. DOI: 10.1103/PhysRevB.99.115433.

[102] X. Song, S. Gong, and Z. Xu, "Propagation of a few-cycle laser pulse in a V-type three-level system", *Opt. Spectrosc.*, vol. 99, no. 4, pp. 517–521, Oct. 2005. DOI: 10.1134/1.2113361.

[103] M. Riesch, P. Tzenov, and C. Jirauschek, "Dynamic simulations of quantum cascade lasers beyond the rotating wave approximation", in *2018 2nd URSI Atlantic Radio Science Meeting (AT-RASC)*, Gran Canaria, Spain: IEEE, Piscataway, NJ, May 2018. DOI: 10.23919/URSI-AT-RASC.2018.8471596.

[104] M. Riesch, J. H. Abundis-Patino, P. Tzenov, and C. Jirauschek, "Efficient simulation of the quantum cascade laser dynamics beyond the rotating wave approximation", in *Proceedings of the International Quantum Cascade Laser School and Workshop (IQCLSW) 2018*, Cassis, France, Sep. 2018.

[105] B. Bidégaray, A. Bourgeade, and D. Reignier, "Introducing physical relaxation terms in Bloch equations", *J. Comput. Phys.*, vol. 170, no. 2, pp. 603–613, Jun. 2001. DOI: 10.1006/jcph.2001.6752.

[106] B. Bidégaray, "Time discretizations for Maxwell-Bloch equations", *Numer. Methods Partial Differ. Equ.*, vol. 19, no. 3, pp. 284–300, Feb. 2003. DOI: 10.1002/num.10046.

[107] O. Saut and A. Bourgeade, "Numerical methods for the bidimensional Maxwell–Bloch equations in nonlinear crystals", *J. Comput. Phys.*, vol. 213, no. 2, pp. 823–843, Apr. 2006. DOI: 10.1016/j.jcp.2005.09.003.

[108] A. Deinega and T. Seideman, "Self-interaction-free approaches for self-consistent solution of the Maxwell-Liouville equations", *Phys. Rev. A*, vol. 89, no. 2, p. 022 501, Feb. 2014. DOI: 10.1103/PhysRevA.89.022501.

[109] M. Riesch, N. Tchipev, S. Senninger, H.-J. Bungartz, and C. Jirauschek, "Performance evaluation of numerical methods for the Maxwell–Liouville–von Neumann equations", *Opt. Quant. Electron.*, vol. 50, no. 2, p. 112, Feb. 2018. DOI: 10.1007/s11082-018-1377-4.

[110] M. Riesch and C. Jirauschek, "Completely positive trace preserving numerical methods for long-term generalized Maxwell-Bloch simulations", in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), 2019 Conference on*, Munich, Germany: IEEE, Piscataway, NJ, Jun. 2019. DOI: 10.1109/CLEOE-EQEC.2019.8873263.

[111] M. Riesch, A. Pikl, and C. Jirauschek, "Completely positive trace preserving methods for the Lindblad equation", in *Numerical Simulation of Optoelectronic Devices (NUSOD), 2020 International Conference on*, Turin, Italy: IEEE, Piscataway, NJ, Sep. 2020. DOI: 10.1109/NUSOD49422.2020.9217670.

[112] Kintechlab, *Electromagnetic template library*, http://fdtd.kintechlab.com/en/start, 2018.

[113] J. Javaloyes and S. Balle, *Freetwm: A simulation tool for semiconductor lasers*, https://onl.uib.eu/Softwares/Freetwm/, 2018.

[114] A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, "Meep: A flexible free-software package for electromagnetic simulations by the FDTD method", *Comput. Phys. Commun.*, vol. 181, no. 3, pp. 687–702, Mar. 2010. DOI: 10.1016/j.cpc.2009.11.008.

[115] M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Solving the Maxwell-Bloch equations efficiently on parallel architectures", in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), 2017 Conference on*, Munich, Germany: IEEE, Piscataway, NJ, Jun. 2017. DOI: 10.1109/CLEOE-EQEC.2017.8087734.

[116] M. Riesch, T. D. Nguyen, and C. Jirauschek, "bertha: Project skeleton for scientific software", *PLOS ONE*, vol. 15, no. 3, e0230557, Mar. 2020. DOI: 10.1371/journal.pone.0230557.

[117] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. Artech House, Boston, 2005.

[118] *mbsolve documentation on GitHub Pages*, https://mriesch-tum.github.io/mbsolve, 2020.

[119] M. Riesch and C. Jirauschek, *mbsolve: An open-source solver tool for the Maxwell-Bloch equations*, https://github.com/mriesch-tum/mbsolve, Jul. 2017.

[120] 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", *Nature Photon.*, vol. 8, no. 6, pp. 462–467, May 2014. DOI: 10.1038/nphoton.2014.85.

[121] M. Singleton, P. Jouy, M. Beck, and J. Faist, "Evidence of linear chirp in mid-infrared quantum cascade lasers", *Optica*, vol. 5, no. 8, pp. 948–953, Aug. 2018. DOI: 10.1364/OPTICA.5.000948.

[122] H. Risken and K. Nummedal, "Self-pulsing in lasers", *J. Appl. Phys.*, vol. 39, no. 10, pp. 4662–4672, Sep. 1968. DOI: 10.1063/1.1655817.

[123] A. Nowogrodzki, "How to support open-source software and stay sane", *Nature*, vol. 571, no. 7763, pp. 133–134, Jul. 2019. DOI: 10.1038/d41586-019-02046-0.

[124] A. Hunt and H. Thomas, *The Pragmatic Programmer: From Journeyman to Master*, 1st ed. Addison-Wesley, Boston, 1999.

[125] A. Prlić and J. B. Procter, "Ten simple rules for the open development of scientific software", *PLoS Comput. Biol.*, vol. 8, no. 12, e1002802, Dec. 2012. DOI: 10.1371/journal.pcbi.1002802.

[126] W. Bangerth and T. Heister, "What makes computational open source software libraries successful?", *Comp. Sci. Disc.*, vol. 6, p. 015 010, Nov. 2013. DOI: 10.1088/1749-4699/6/1/015010.

[127] G. Wilson, D. A. Aruliah, C. T. Brown, N. P. Chue Hong, M. Davis, R. T. Guy, S. H. D. Haddock, K. D. Huff, I. M. Mitchell, M. D. Plumbley, B. Waugh, E. P. White, and P. Wilson, "Best practices for scientific computing", *PLoS Biol.*, vol. 12, no. 1, e1001745, Jan. 2014. DOI: 10.1371/journal.pbio.1001745.

[128] G. Wilson, J. Bryan, K. Cranston, J. Kitzes, L. Nederbragt, and T. K. Teal, "Good enough practices in scientific computing", *PLoS Comput. Biol.*, vol. 13, no. 6, e1005510, Jun. 2017. DOI: 10.1371/journal.pcbi.1005510.

[129] T. Schlauch, M. Meinel, and C. Haupt, *DLR software engineering guidelines*, https://doi.org/10.5281/zenodo.1344612, version 1.0.0, Aug. 2018. DOI: 10.5281/zenodo.1344612.

[130] Netherlands eScience Center, *Software development guide*, https://guide.esciencecenter.nl, Aug. 2019.

[131] *Clean C++ project for you to use*, https://github.com/kracejic/cleanCppProject, Oct. 2015.

[132] A. Ioannides, *Python package template project for kick-starting new Python projects.* https://github.com/AlexIoannides/py-package-template, Nov. 2018.

[133] J. Highsmith, *Agile Project Management: Creating Innovative Products*, 2nd ed., A. Cockburn and J. Highsmith, Eds. Addison-Wesley, Boston, 2010.

[134] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.*, "The FAIR guiding principles for scientific data management and stewardship", *Sci. Data*, vol. 3, p. 160 018, Mar. 2016. DOI: 10.1038/sdata.2016.18.

[135] A.-L. Lamprecht, L. Garcia, M. Kuzak, C. Martinez, R. Arcila, E. Martin Del Pico, V. Dominguez Del Angel, S. van de Sandt, J. Ison, P. A. Martinez, *et al.*, "Towards FAIR principles for research software", *Data Sci.*, vol. 3, no. 1, pp. 37–59, Jun. 2019. DOI: 10.3233/DS-190026.

[136] M. Riesch and C. Jirauschek, *bertha: Project skeleton for scientific software (C++ with Python interface)*, https://gitlab.com/cph-tum/bertha, Dec. 2019.

[137] *Conda feedstock for bertha*, https://github.com/conda-forge/bertha-feedstock, Dec. 2019.

[138] GitLab Inc., *Introduction to GitLab Flow*, https://docs.gitlab.com/ee/topics/gitlab_flow.html, 2019.

[139] M. Riesch and C. Jirauschek, *bertha: Project skeleton for scientific software (C++ with Python interface)*, https://github.com/cph-tum/bertha, Dec. 2019.

[140] M. Riesch, *Add template for small to medium scientific software projects.* https://gitlab.com/gitlab-org/gitlab/issues/35293, Dec. 2019.

[141] *Bertha documentation on GitLab Pages*, https://cph-tum.gitlab.io/bertha, Dec. 2019.

[142] M. Riesch, M. Haider, and C. Jirauschek, "Project skeletons for scientific software", in *Numerical Simulation of Optoelectronic Devices (NUSOD), 2020 International Conference on*, Turin, Italy: IEEE, Piscataway, NJ, Sep. 2020. DOI: 10.1109/NUSOD49422.2020.9217756.

[143] M. Haider, M. Riesch, and C. Jirauschek, "Realization of best practices in software engineering and scientific writing through ready-to-use project skeletons", *Opt. Quant. Electron.*, vol. 53, p. 568, Oct. 2021. DOI: 10.1007/s11082-021-03192-4.

[144] M. Riesch and C. Jirauschek, "Analyzing the positivity preservation of numerical methods for the Liouville-von Neumann equation", *J. Comput. Phys.*, vol. 390, pp. 290–296, Feb. 2019. DOI: 10.1016/j.jcp.2019.04.006.

[145] K. Yee, "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media", *IEEE Trans. Antennas. Propag.*, vol. 14, no. 3, pp. 302–307, May 1966. DOI: 10.1109/TAP.1966.1138693.

[146] M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures", in *Proceedings of the Platform for Advanced Scientific Computing Conference*, Zurich, Switzerland: ACM, New York, NY, Jun. 2019, 5:1–5:8. DOI: 10.1145/3324989.3325715.

[147] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I*, 2nd ed. Springer-Verlag, Berlin Heidelberg, 1993.

[148] B. Hellsing and H. Metiu, "An efficient method for solving the quantum Liouville equation: Applications to electronic absorption spectroscopy", *Chem. Phys. Lett.*, vol. 127, no. 1, pp. 45–49, May 1986. DOI: 10.1016/S0009-2614(86)80206-8.

[149] M. Berman, R. Kosloff, and H. Tal-Ezer, "Solution of the time-dependent Liouville-von Neumann equation: Dissipative evolution", *J. Phys. A: Math. Gen.*, vol. 25, no. 5, pp. 1283–1307, Mar. 1992. DOI: 10.1088/0305-4470/25/5/031.

[150] M. E. Songolo and B. Bidégaray-Fesquet, "Nonstandard finite-difference schemes for the two-level Bloch model", *Int. J. Model. Simul. Sci. Comput.*, vol. 9, no. 4, p. 1 850 033, Apr. 2018. DOI: 10.1142/S1793962318500332.

[151] G. Guennebaud, B. Jacob, *et al.*, *Eigen v3*, http://eigen.tuxfamily.org, 2010.

[152] C. Moler and C. V. Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later", *SIAM Rev.*, vol. 45, no. 1, pp. 3–49, 2003. DOI: 10.1137/S00361445024180.

[153] A. H. Al-Mohy and N. J. Higham, "Computing the action of the matrix exponential, with an application to exponential integrators", *SIAM J. Sci. Comput.*, vol. 33, no. 2, pp. 488–511, Mar. 2011. DOI: 10.1137/100788860.

[154] M. Hochbruck and C. Lubich, "On Krylov subspace approximations to the matrix exponential operator", *SIAM J. Numer. Anal.*, vol. 34, no. 5, pp. 1911–1925, Oct. 1997. DOI: 10.1137/S0036142995280572.

[155] M. Hochbruck, C. Lubich, and H. Selhofer, "Exponential integrators for large systems of differential equations", *SIAM J. Sci. Comput.*, vol. 19, no. 5, pp. 1552–1574, Sep. 1998. DOI: 10.1137/S1064827595295337.

[156] M. Pototschnig, J. Niegemann, L. Tkeshelashvili, and K. Busch, "Time-domain simulations of the nonlinear Maxwell equations using operator-exponential methods", *IEEE Trans. Antennas Propag.*, vol. 57, no. 2, pp. 475–483, Feb. 2009. DOI: 10.1109/TAP.2008.2011181.

[157] L. Guduff, A. J. Allami, C. van Heijenoort, J.-N. Dumez, and I. Kuprov, "Efficient simulation of ultrafast magnetic resonance experiments", *Phys. Chem. Chem. Phys.*, vol. 19, no. 27, pp. 17 577–17 586, Jun. 2017. DOI: 10.1039/C7CP03074F.

[158] R. Kosloff, "Propagation methods for quantum molecular dynamics", *Annu. Rev. Phys. Chem.*, vol. 45, no. 1, pp. 145–178, Oct. 1994. DOI: 10.1146/annurev.pc.45.100194.001045.

[159] D. Basilewitsch, L. Marder, and C. P. Koch, "Dissipative quantum dynamics and optimal control using iterative time ordering: An application to superconducting qubits", *Eur. Phys. J. B*, vol. 91, no. 7, p. 161, Jul. 2018. DOI: 10.1140/epjb/e2018-90224-4.

[160] G. Strang, "On the construction and comparison of difference schemes", *SIAM J. Numer. Anal.*, vol. 5, no. 3, pp. 506–517, Sep. 1968. DOI: 10.1137/0705041.

[161] D. Coppersmith and S. Winograd, "Matrix multiplication via arithmetic progressions", *J. Symb. Comput.*, vol. 9, no. 3, pp. 251–280, Mar. 1990. DOI: 10.1016/S0747-7171(08)80013-2.

[162] R. Courant and D. Hilbert, *Methods of Mathematical Physics*. Wiley-VCH, Weinheim, 2004, vol. 1.

[163] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. The MIT Press, Cambridge, MA, 2009.

[164] M. Wu, S. Chen, K. J. Schafer, and M. B. Gaarde, "Ultrafast time-dependent absorption in a macroscopic three-level helium gas", *Phys. Rev. A*, vol. 87, no. 1, p. 013 828, Jan. 2013. DOI: 10.1103/PhysRevA.87.013828.

[165] T. F. Havel, "Robust procedures for converting among Lindblad, Kraus and matrix representations of quantum dynamical semigroups", *J. Math. Phys.*, vol. 44, no. 2, pp. 534–557, Jan. 2003. DOI: 10.1063/1.1518555.

[166] M. Riesch, *Supplementary MATLAB code*, https://doi.org/10.5281/zenodo.2560305, Feb. 2019.

[167] M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Numerical methods for the Maxwell-Liouville-von Neumann (MLN) equations", in *2018 2nd URSI Atlantic Radio Science Meeting (AT-RASC)*, Gran Canaria, Spain: IEEE, Piscataway, NJ, May 2018. DOI: 10.23919/URSI-AT-RASC.2018.8471589.

[168] J. Gallier and D. Xu, "Computing exponentials of skew symmetric matrices and logarithms of orthogonal matrices", *Int. J. Robot. Autom.*, vol. 18, no. 1, pp. 10–20, 2003.

[169] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge University Press, Cambridge, 1990.

[170] G. Demeter, "Solving the Maxwell-Bloch equations for resonant nonlinear optics using GPUs", *Comput. Phys. Commun.*, vol. 184, no. 4, pp. 1203–1210, Apr. 2013. DOI: 10.1016/j.cpc.2012.12.019.

[171] S. Krishnamoorthy, M. Baskaran, U. Bondhugula, J. Ramanujam, A. Rountev, and P. Sadayappan, "Effective automatic parallelization of stencil computations", *SIGPLAN Not.*, vol. 42, no. 6, pp. 235–244, Jun. 2007. DOI: 10.1145/1273442.1250761.

[172] J. Treibig, G. Hager, and G. Wellein, "Likwid: A lightweight performance-oriented tool suite for x86 multicore environments", in *Proceedings of PSTI2010, the First International Workshop on Parallel Software Tools and Tool Infrastructures*, San Diego, CA: IEEE, Piscataway, NJ, Sep. 2010, pp. 207–216. DOI: 10.1109/ICPPW.2010.38.

[173] F. P. Mezzapesa, K. Garrasi, J. Schmidt, L. Salemi, V. Pistore, L. Li, A. G. Davies, E. H. Linfield, M. Riesch, C. Jirauschek, T. Carey, F. Torrisi, A. C. Ferrari, and M. S. Vitiello, "Terahertz frequency combs exploiting an on-chip, solution processed, graphene-quantum cascade laser coupled-cavity", *ACS Photonics*, vol. 7, no. 12, pp. 3489–3498, Dec. 2020. DOI: 10.1021/acsphotonics.0c01523.

[174] R. Paiella, F. Capasso, C. Gmachl, D. L. Sivco, J. N. Baillargeon, A. L. Hutchinson, A. Y. Cho, and H. Liu, "Self-mode-locking of quantum cascade lasers with giant ultrafast optical nonlinearities", *Science*, vol. 290, no. 5497, pp. 1739–1742, Dec. 2000. DOI: 10.1126/science.290.5497.1739.

[175] A. Soibel, F. Capasso, C. Gmachl, M. Peabody, A. Sergent, R. Paiella, D. Sivco, A. Cho, and H. Liu, "Stability of pulse emission and enhancement of intracavity second-harmonic generation in self-mode-locked quantum cascade lasers", *IEEE J. Quantum Electron.*, vol. 40, no. 3, pp. 197–204, Mar. 2004. DOI: 10.1109/JQE.2003.823036.

[176] C. Y. Wang, L. Kuznetsova, V.-M. Gkortsas, L. Diehl, F. X. Kärtner, M. A. Belkin, A. Belyanin, X. Li, D. Ham, H. Schneider, P. Grant, C. Y. Song, S. Haffouz, Z. R. Wasilewski, H.-C. Liu, and F. Capasso, "Mode-locked pulses from mid-infrared quantum cascade lasers", *Opt. Express*, vol. 17, no. 15, pp. 12 929–12 943, Jul. 2009. DOI: 10.1364/OE.17.012929.

[177] F. Bloch and A. Siegert, "Magnetic resonance for nonrotating fields", *Phys. Rev.*, vol. 57, no. 6, p. 522, Mar. 1940. DOI: 10.1103/PhysRev.57.522.

[178] D. Kazakov, M. Piccardo, Y. Wang, P. Chevalier, T. S. Mansuripur, F. Xie, C.-e. Zah, K. Lascola, A. Belyanin, and F. Capasso, "Self-starting harmonic frequency comb generation in a quantum cascade laser", *Nat. Photonics*, vol. 11, no. 12, p. 789, Oct. 2017. DOI: 10.1038/s41566-017-0026-y.

[179] H. Li, P. Laffaille, D. Gacemi, M. Apfel, C. Sirtori, J. Leonardon, G. Santarelli, M. Rösch, G. Scalari, M. Beck, J. Faist, W. Hänsel, R. Holzwarth, and S. Barbieri, "Dynamics of ultra-broadband terahertz quantum cascade lasers for comb operation", *Opt. Express*, vol. 23, no. 26, pp. 33 270–33 294, Dec. 2015. DOI: 10.1364/OE.23.033270.

[180] P. Gellie, S. Barbieri, J.-F. Lampin, P. Filloux, C. Manquest, C. Sirtori, I. Sagnes, S. P. Khanna, E. H. Linfield, A. G. Davies, H. Beere, and D. Ritchie, "Injection-locking of terahertz quantum cascade lasers up to 35GHz using RF amplitude modulation", *Opt. Express*, vol. 18, no. 20, pp. 20 799–20 816, Sep. 2010. DOI: 10.1364/OE.18.020799.

[181] A. Mottaghizadeh, D. Gacemi, P. Laffaille, H. Li, M. Amanti, C. Sirtori, G. Santarelli, W. Hänsel, R. Holzwart, L. H. Li, E. H. Linfield, and S. Barbieri, "5-ps-long terahertz pulses from an active-mode-locked quantum cascade laser", *Optica*, vol. 4, no. 1, pp. 168–171, Jan. 2017. DOI: 10.1364/OPTICA.4.000168.

[182] F. X. Kärtner, C. Jirauschek, F. Capasso, A. Soibel, A. Belyanin, and J. Faist, "Mode locking of quantum cascade lasers (theory)", Workshop in Infrared Physics, Seville, Spain, Jan. 2004.

[183] S. Kohen, B. S. Williams, and Q. Hu, "Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators", *J. Appl. Phys.*, vol. 97, no. 5, p. 053 106, Feb. 2005. DOI: 10.1063/1.1855394.

[184] W. Maineult, L. Ding, P. Gellie, P. Filloux, C. Sirtori, S. Barbieri, T. Akalin, J.-F. Lampin, I. Sagnes, H. E. Beere, and D. A. Ritchie, "Microwave modulation of terahertz quantum cascade lasers: A transmission-line approach", *Appl. Phys. Lett.*, vol. 96, no. 2, p. 021 108, Jan. 2010. DOI: 10.1063/1.3284518.

[185] P. Tzenov, D. Burghoff, M. Riesch, Q. Hu, and C. Jirauschek, "Coupled transmission line/Maxwell-Bloch simulation approach for analysis of active mode locking in terahertz quantum cascade lasers", in *Proceedings of the International Quantum Cascade Lasers School and Workshop (IQCLSW) 2016*, Cambridge, United Kingdom, Sep. 2016.

[186] P. Tzenov, "Modeling and simulations of quantum cascade lasers for frequency comb generation", PhD thesis, Technische Universität München, Sep. 2018.

[187] M. Piccardo, D. Kazakov, N. A. Rubin, P. Chevalier, Y. Wang, F. Xie, K. Lascola, A. Belyanin, and F. Capasso, "Time-dependent population inversion gratings in laser frequency combs", *Optica*, vol. 5, no. 4, pp. 475–478, Apr. 2018. DOI: 10.1364/OPTICA.5.000475.

[188] V. Pistore, "Modelocking of THz quantum cascade lasers: Dispersion control and non-linearities", PhD thesis, Sorbonne Université, Nov. 2019.

[189] F. Wang, V. Pistore, M. Riesch, H. Nong, P.-B. Vigneron, R. Colombelli, O. Parillaud, J. Mangeney, J. Tignon, C. Jirauschek, and S. S. Dhillon, *Experimental and simulation results of "Ultrafast response of harmonic modelocked THz laser"*, Apr. 2020. DOI: 10.5281/zenodo.3689614.

[190] V. Bianchi, T. Carey, L. Viti, L. Li, E. H. Linfield, A. G. Davies, A. Tredicucci, D. Yoon, P. G. Karagiannidis, L. Lombardi, F. Tomarchio, A. C. Ferrari, F. Torrisi, and M. S. Vitiello, "Terahertz saturable absorbers from liquid phase exfoliation of graphite", *Nat. Commun.*, vol. 8, no. 1, pp. 1–9, Jun. 2017. DOI: 10.1038/ncomms15763.

[191] Q. Bao, H. Zhang, Y. Wang, Z. Ni, Y. Yan, Z. X. Shen, K. P. Loh, and D. Y. Tang, "Atomic-layer graphene as a saturable absorber for ultrafast pulsed lasers", *Adv. Funct. Mater.*, vol. 19, no. 19, pp. 3077–3083, Oct. 2009. DOI: 10.1002/adfm.200901007.

[192] Y. Yang, A. Paulsen, D. Burghoff, J. L. Reno, and Q. Hu, "Lateral heterogeneous integration of quantum cascade lasers", *ACS Photonics*, vol. 5, no. 7, pp. 2742–2747, Jun. 2018. DOI: 10.1021/acsphotonics.8b00507.

[193] C. Jirauschek and M. Haider, "Traveling wave effects in microwave quantum photonics", in *2021 International Applied Computational Electromagnetics Society Symposium (ACES)*, Hamilton, Canada: IEEE, Piscataway, NJ, Sep. 2021, pp. 1–3. DOI: 10.1109/ACES53325.2021.00092.

[194] A. Forrer, Y. Wang, M. Beck, A. Belyanin, J. Faist, and G. Scalari, "Self-starting harmonic comb emission in THz quantum cascade lasers", *Appl. Phys. Lett.*, vol. 118, no. 13, p. 131 112, Mar. 2021. DOI: 10.1063/5.0041339.

[195] T. M. Fischer, "On the algorithm by Al-Mohy and Higham for computing the action of the matrix exponential: A posteriori roundoff error estimation", *Linear Algebra Its Appl.*, vol. 531, pp. 141–168, Oct. 2017. DOI: 10.1016/j.laa.2017.05.042.

[196] H. Choi, J.-W. Baek, and K.-Y. Jung, "Comprehensive study on numerical aspects of modified Lorentz model-based dispersive FDTD formulations", *IEEE Trans. Antennas Propag.*, vol. 67, no. 12, pp. 7643–7648, Aug. 2019. DOI: 10.1109/TAP.2019.2934779.

[197] L. Seitner, J. Popp, M. Riesch, M. Haider, and C. Jirauschek, "Group velocity dispersion in terahertz frequency combs within a generalized Maxwell-Bloch framework", *J. Phys. Conf. Ser.*, vol. 2090, no. 1, p. 012 082, Nov. 2021. DOI: 10.1088/1742-6596/2090/1/012082.

[198] A. Bismuto, R. Terazzi, B. Hinkov, M. Beck, and J. Faist, "Fully automatized quantum cascade laser design by genetic optimization", *Appl. Phys. Lett.*, vol. 101, no. 2, p. 021 103, Jul. 2012. DOI: 10.1063/1.4734389.

[199] D. Mueller and G. Triplett, "Development of a multi-objective evolutionary algorithm for strain-enhanced quantum cascade lasers", *Photonics*, vol. 3, no. 3, p. 44, Jul. 2016. DOI: 10.3390/photonics3030044.

[200] S. Hughes, "Breakdown of the area theorem: Carrier-wave Rabi flopping of femtosecond optical pulses", *Phys. Rev. Lett.*, vol. 81, pp. 3363–3366, 16 Oct. 1998. DOI: 10.1103/PhysRevLett.81.3363.

Publications

Journal articles

- M. Riesch, N. Tchipev, S. Senninger, H.-J. Bungartz, and C. Jirauschek, "Performance evaluation of numerical methods for the Maxwell–Liouville–von Neumann equations", *Opt. Quant. Electron.*, vol. 50, no. 2, p. 112, Feb. 2018. DOI: 10.1007/s11082-018-1377-4.
- M. Riesch and C. Jirauschek, "Analyzing the positivity preservation of numerical methods for the Liouville-von Neumann equation", *J. Comput. Phys.*, vol. 390, pp. 290–296, Feb. 2019. DOI: 10.1016/j.jcp.2019.04.006.
- C. Jirauschek, M. Riesch, and P. Tzenov, "Optoelectronic device simulations based on macroscopic Maxwell–Bloch equations", *Adv. Theor. Simul.*, vol. 2, no. 8, p. 1 900 018, Aug. 2019. DOI: 10.1002/adts.201900018.
- M. Riesch, T. D. Nguyen, and C. Jirauschek, "bertha: Project skeleton for scientific software", *PLOS ONE*, vol. 15, no. 3, e0230557, Mar. 2020. DOI: 10.1371/journal.pone.0230557.
- F. Wang, V. Pistore, M. Riesch, H. Nong, P.-B. Vigneron, R. Colombelli, O. Parillaud, J. Mangeney, J. Tignon, C. Jirauschek, and S. S. Dhillon, "Ultrafast response of active and self-starting harmonic modelocked THz laser", *Light Sci. Appl.*, vol. 9, no. 1, p. 51, Apr. 2020. DOI: 10.1038/s41377-020-0288-x.
- F. P. Mezzapesa, K. Garrasi, J. Schmidt, L. Salemi, V. Pistore, L. Li, A. G. Davies, E. H. Linfield, M. Riesch, C. Jirauschek, T. Carey, F. Torrisi, A. C. Ferrari, and M. S. Vitiello, "Terahertz frequency combs exploiting an on-chip, solution processed, graphene-quantum cascade laser coupled-cavity", *ACS Photonics*, vol. 7, no. 12, pp. 3489–3498, Dec. 2020. DOI: 10.1021/acsphotonics.0c01523.
- M. Haider, M. Riesch, and C. Jirauschek, "Realization of best practices in software engineering and scientific writing through ready-to-use project skeletons", *Opt. Quant. Electron.*, vol. 53, p. 568, Oct. 2021. DOI: 10.1007/s11082-021-03192-4.
- M. Riesch and C. Jirauschek, "mbsolve: An open-source solver tool for the Maxwell-Bloch equations", *Comput. Phys. Commun.*, vol. 268, p. 108 097, Nov. 2021. DOI: 10.1016/j.cpc.2021.108097.

Contributions to conferences

- P. Tzenov, D. Burghoff, M. Riesch, Q. Hu, and C. Jirauschek, "Coupled transmission line/Maxwell-Bloch simulation approach for analysis of active mode locking in terahertz quantum cascade lasers", in *Proceedings of the International Quantum Cascade Lasers School and Workshop (IQCLSW) 2016*, Cambridge, United Kingdom, Sep. 2016.
- M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Solving the Maxwell-Bloch equations efficiently on parallel architectures", in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), 2017 Conference on*, Munich, Germany: IEEE, Piscataway, NJ, Jun. 2017. DOI: 10.1109/CLEOE-EQEC.2017.8087734.

- M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Performance evaluation of numerical methods for the Maxwell-Liouville equations", in *Numerical Simulation of Optoelectronic Devices (NUSOD), 2017 International Conference on*, Copenhagen, Denmark: IEEE, Piscataway, NJ, Jul. 2017, pp. 223–224. DOI: 10.1109/NUSOD.2017.8010072.

- M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Numerical methods for the Maxwell-Liouville-von Neumann (MLN) equations", in *2018 2nd URSI Atlantic Radio Science Meeting (AT-RASC)*, Gran Canaria, Spain: IEEE, Piscataway, NJ, May 2018. DOI: 10.23919/URSI-AT-RASC.2018.8471589.

- M. Riesch, P. Tzenov, and C. Jirauschek, "Dynamic simulations of quantum cascade lasers beyond the rotating wave approximation", in *2018 2nd URSI Atlantic Radio Science Meeting (AT-RASC)*, Gran Canaria, Spain: IEEE, Piscataway, NJ, May 2018. DOI: 10.23919/URSI-AT-RASC.2018.8471596.

- M. Riesch, J. H. Abundis-Patino, P. Tzenov, and C. Jirauschek, "Efficient simulation of the quantum cascade laser dynamics beyond the rotating wave approximation", in *Proceedings of the International Quantum Cascade Laser School and Workshop (IQCLSW) 2018*, Cassis, France, Sep. 2018.

- J. H. Abundis-Patino, M. Riesch, P. Tzenov, and C. Jirauschek, "Colliding pulse mode locking of quantum cascade lasers", in *Ultrafast Phenomena 2018 (UP 2018), XXI International Conference on*, Hamburg, Germany: EDP Sciences, Les Ulis, France, Apr. 2019. DOI: 10.1051/epjconf/201920501018.

- M. Riesch and C. Jirauschek, "Completely positive trace preserving numerical methods for long-term generalized Maxwell-Bloch simulations", in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), 2019 Conference on*, Munich, Germany: IEEE, Piscataway, NJ, Jun. 2019. DOI: 10.1109/CLEOE-EQEC.2019.8873263.

- M. Riesch, V. Pistore, F. Wang, S. S. Dhillon, and C. Jirauschek, "Modeling self-starting harmonic mode locking in terahertz quantum cascade lasers", in *Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC), 2019 Conference on*, Munich, Germany: IEEE, Piscataway, NJ, Jun. 2019. DOI: 10.1109/CLEOE-EQEC.2019.8872672.

- M. Riesch, N. Tchipev, H.-J. Bungartz, and C. Jirauschek, "Numerical simulation of the quantum cascade laser dynamics on parallel architectures", in *Proceedings of the Platform for Advanced Scientific Computing Conference*, Zurich, Switzerland: ACM, New York, NY, Jun. 2019, 5:1–5:8. DOI: 10.1145/3324989.3325715.

- V. Pistore, F. Wang, M. Riesch, H. Nong, P.-B. Vigneron, R. Colombelli, O. Parillaud, C. Jirauschek, J. Mangeney, J. Tignon, and S. S. Dhillon, "Self-starting harmonic emission and active harmonic modelocking in THz QCLs", in *Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2019, 44th International Conference on*, Paris, France: IEEE, Piscataway, NJ, Sep. 2019. DOI: 10.1109/IRMMW-THz.2019.8873879.

- V. Pistore, F. Wang, M. Riesch, H. Nong, P.-B. Vigneron, R. Colombelli, O. Parillaud, C. Jirauschek, J. Mangeney, J. Tignon, and S. S. Dhillon, "Active harmonic modelocking and self-starting harmonic emission in THz QCLs", in *Proceedings of the Infrared Terahertz Quantum Workshop (ITQW) 2019*, Ojai, CA, USA, Sep. 2019.

- M. Riesch and C. Jirauschek, "Open-source simulation software for quantum cascade lasers", in *Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2019, 44th International Conference on*, Paris, France: IEEE, Piscataway, NJ, Sep. 2019. DOI: 10.1109/IRMMW-THz.2019.8874442.

- M. Riesch and C. Jirauschek, "Modeling and simulation of mode locking in quantum cascade lasers", in *Proceedings of the Infrared Terahertz Quantum Workshop (ITQW) 2019*, Ojai, CA, USA, Sep. 2019.
- M. Riesch, M. Haider, and C. Jirauschek, "Project skeletons for scientific software", in *Numerical Simulation of Optoelectronic Devices (NUSOD), 2020 International Conference on*, Turin, Italy: IEEE, Piscataway, NJ, Sep. 2020. DOI: 10.1109/NUSOD49422.2020.9217756.
- M. Riesch, A. Pikl, and C. Jirauschek, "Completely positive trace preserving methods for the Lindblad equation", in *Numerical Simulation of Optoelectronic Devices (NUSOD), 2020 International Conference on*, Turin, Italy: IEEE, Piscataway, NJ, Sep. 2020. DOI: 10.1109/NUSOD49422.2020.9217670.
- L. Seitner, J. Popp, M. Riesch, M. Haider, and C. Jirauschek, "Group velocity dispersion in terahertz frequency combs within a generalized Maxwell-Bloch framework", *J. Phys. Conf. Ser.*, vol. 2090, no. 1, p. 012 082, Nov. 2021. DOI: 10.1088/1742-6596/2090/1/012082.

Acknowledgments

Ladies and gentlemen, and friends of my daughter. There comes a time in every wedding reception when the man who paid the for the damn thing is allowed to speak a word or two of his own. And I should like to take this opportunity, sloshed as I may be, to say a word or two...

— *Rowan Atkinson, Father of the Bride Speech*

And in the end the love you take
is equal to the love you make.

— *The Beatles, The End*

At this point I would like to express my deepest gratitude to many people who have supported me during the elaboration of this thesis.

My advisor Christian Jirauschek offered me the possibility to elaborate this interesting topic in the form of a Master's as well as a Ph.D. thesis, respectively. I appreciated very much the informal and productive atmosphere that he established in the Computational Photonics Group. He always had time to answer my questions and was open for new ideas, even in those cases where I did not exactly follow the pragmatic approach.

Petar Tzenov was the advisor of my Master's thesis and later became my colleague when I started my doctoral degree. Week after week we had inspiring discussions which gave my a very good start into the field of quantum cascade lasers. It was great to work with Petar also on a personal level, and with amusement I look back to the first conference we attended together.

A similar thing can be said about my office mates Jesus Humberto Abundis Patino, Tim Albes, Johannes Popp and Mark Schmidt, as well as my colleagues Michael Haider and Michael Rinderle from the other groups. Apart from research-related issues we also found completely different things to discuss, often in the scope of the (infamous) International Conference on Mostly Scientific Topics (ICMST).

The staff of the Chair of Nano and Quantum Sensors was always very supportive in organizational matters. I would like to thank Markus Becherer, Katrin Blahetek, Susanne Maier, Rosi Mittermeier and Lucia Weik for their help when I had to fight bureaucratic hurdles.

Over the years I had the pleasure to serve as thesis advisor of several students who did an excellent job. I would like to thank Mariem Kthiri, Tien Dat Nguyen, Alek Pikl, Sebastian Senninger, Wenhua Shi, Christian Widmann, Yi Zhang and Yiming Zhao for their contributions to the mbsolve project.

The conception of the bertha skeleton project did benefit significantly from the exchange with fellow researchers. I gratefully acknowledge the stimulating discussions with Wolfgang Bangerth, Richard Barnes, Carina Haupt, Alexander Valavanis and Joachim Wuttke, as well as the support of Isuru Fernando and Ray Donnelly during the implementation of the conda recipe.

It was always a pleasure to collaborate with other research groups. Nikola Tchipev and Hans-Joachim Bungartz from the Chair of Scientific Computing in Computer Science at the Technical University of Munich provided excellent advice during the implementation and optimization of the parallelized solvers. Then, we had a stimulating exchange of ideas with Valentino Pistore and Sukhdeep Dhillon from the Laboratoire de Physique de l'École Normale Supérieure in Paris. Valentino later went on to pursue a post-doc position at the Istituto Nanoscienze Consiglio Nazionale delle Ricerche in Pisa in the group of Miriam Vitiello, with whom we had a very productive collaboration.

Valentino deserves special mention here as he also was always up for interesting discussions, even at a later stage of the evening. Those discussions usually included one or two beers and not only spanned

research topics but also questions about life and the universe in general, and the history of Western Europe between the 14th and 19th century in particular.

Last, but not least I would like to thank my friends and my family, who have always been there for me. They have brought change into my life and reminded me from time to time that other things apart from the thesis are important too. That said, I am not sure whether I would have chosen to pursue a doctoral degree without my brother Christian as a role model. Over all those years the chats we had about electronics and computer science have served as constant source of motivation for me.

Finally, and especially, I would like to thank my parents Doris and Herbert as well as my lovely wife Andrea for their support, their patience, and their encouragement!

Curriculum vitae

Out of all the things I have lost, I miss my mind the most.
— attributed to Mark Twain, Ozzy Osbourne, ...

Personal data

Michael Riesch
Born October 26, 1988 in Bludenz, Austria

Education

1995 – 2003	Primary and secondary school
Bürs, Austria	
2003 – 2008	Höhere Technische Bundeslehr- und Versuchsanstalt Rankweil
Rankweil, Austria	*(Higher technical education institute)*
	Matura (Graduation) in 2008 with distinction.

Civilian service

2008 – 2009	Austrian Red Cross
Bludenz, Austria	*Rettungssanitäter (Emergency medical technician – Intermediate)*

Academic courses

2009 – 2012	Technical University of Munich
Munich, Germany	*Bachelor's degree "Electrical Engineering and Information Technology", Bachelor's thesis "Via-hole contacting of PET sheets", "passed with merit"*
2012 – 2015	Technical University of Munich
Munich, Germany	*Master's degree "Electrical Engineering and Information Technology", Master's thesis "A High-Level Control System for an Electro-Mechanical Resuscitation Device", "passed with merit"*
2014	Universidad de Cádiz
Cádiz, Spain	*ERASMUS semester "Ingeniería Informática"*
2013 – 2016	Technical University of Munich
Munich, Germany	*Master's degree "Computational Science and Engineering", Master's thesis "Efficient Simulation of Active Mode-Locking in Quantum Cascade Lasers", "passed with merit"*

Professional experience

2016 – 2020	Technical University of Munich
Munich, Germany	*Research assistant at the Computational Photonics Group*
2020 –	WolfVision GmbH
Klaus i. Vbg., Austria	*Embedded Linux developer*

Good night, Austin, Texas, wherever you are!
— *Frank Zappa*