Development and Optimisation of Modelling Methods and Algorithms for Terahertz Range Radiation Sources Based on Quantum Well Heterostructures

REENO REEDER
Dissertation was accepted for the defence of the degree of Doctor of Philosophy in Engineering on May 12, 2014.

Supervisor: Enn Velmre, Prof. Emeritus, Ph.D. Thomas Johann Seebeck Department of Electronics

Co-Supervisor: Andres Udal, Senior Researcher, Ph.D. Department of Computer Control

Opponents: Albrecht Zur, Prof. Emer., Dr. Habil., UAS, Kiel, Germany
Alvo Aabloo, Prof., Ph.D., University of Tartu, Estonia

Defence of the thesis: May 30, 2014

Declaration: I hereby declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for another doctoral or equivalent academic degree.

Reeno Reeder

Copyright: Reeno Reeder 2014
ISSN 1406-4723
ISBN 978-9949-23-609-1 (publication)
ISBN 978-9949-23-610-7 (PDF)
Kvantaukudega heterostruktuuridel põhinevate terahertskiirgurite modelleerimismeeodite ja -algoritmide arendus ja optimeerimine

REENO REEDER
## Contents

List of publications ............................................. 7

Author’s Contribution to the Publications .................... 8

Introduction ...................................................... 9

List of acronyms .................................................. 11

1 Overview of the terahertz region ............................... 12
   1.1 Overview of technology in the THz region ................. 12
   1.2 THz radiation sources ..................................... 14

2 Time-independent Schrödinger equation numerical solution methods. Application to digitally graded GaAs/AlGaAs parabolic quantum wells ............................ 18
   2.1 Introduction ............................................... 18
   2.2 Choosing a suitable numerical method .................... 19
      2.2.1 Shooting method ...................................... 19
      2.2.2 Energy and wave function coupled solution method (Newton method) .................................. 20
      2.2.3 Method based on matrix eigenvalue standard solvers ...................................................... 22
      2.2.4 Comparison of all three methods ...................... 24
   2.3 Optimising digitally graded potentials .................... 25
      2.3.1 Introduction .......................................... 25
      2.3.2 Building initial potential ............................ 25
      2.3.3 Optimising the potential ............................ 26
   2.4 Conclusion .................................................. 27

3 Lateral transport task of a quantum well-based broadband terahertz emitter .................................... 28
   3.1 Introduction ............................................... 28
   3.2 Shifted Fermi-Dirac distribution function ................ 30
   3.3 Calculation methodology .................................. 31
3.3.1 Overview ........................................... 31
3.3.2 Calculation of raw scattering rates .......... 31
3.3.3 Calculation of quasi-Fermi energies .......... 35
3.3.4 Calculation of mean scattering rates .......... 35
3.3.5 Calculation of subband populations using rate equations 36
3.3.6 Calculation of emission characteristics .... 37
3.3.7 Overall calculation automation ................. 38
3.4 Conclusion ........................................... 38

4 Optimisation of mathematical models for terahertz range quantum cascade lasers 40
4.1 Introduction ........................................... 40
4.2 Improvement of the Schrödinger solver ........ 45
4.3 Improvements due to the reuse of data .......... 48
4.4 Integration improvements .......................... 50
4.5 Conclusion ........................................... 53

Conclusion ............................................. 54

References ............................................. 55

Acknowledgements ...................................... 61

Abstract ............................................... 62

Kokkuvõte ............................................... 63

Appendices ............................................. 65
  Article I ............................................... 67
  Article II ............................................. 73
  Article III ........................................... 91
  Article IV ........................................... 99
  Article V ............................................. 105
  Curriculum vitae (in English) ....................... 109
  Curriculum vitae (in Estonian) ..................... 112

Dissertations defended at Tallinn University of Technology on Informatics and System Engineering 115

6
List of publications

Here is a list of the main publications upon which this work stands, according to Tallinn University of Technology’s classification scheme. Copies of the publications can be found in Appendices I-V.


Author’s Contribution to the Publications

Here is the list of author’s contributions to the articles given in the previous page.

- **I.** Analysis of the possibilities to design parabolic quantum wells using MBE; development of the algorithm to design digitally graded parabolic quantum wells; experimental simulations to get emission characteristics for general and digitally graded parabolic quantum wells: analysis of $\Delta E$ deviation; writing of the article.

- **II.** Efficiency analysis and measurements of timing of well known Schrödinger solution algorithms; review of the article.

- **III.** Analysis of the ways to produce broadband terahertz emission in room temperature; development of the lateral transport model of electrons; simulations of emission characteristics for digitally graded parabolic quantum wells; writing of the article.

- **IV.** The selection and analysis of publications; preparing an overview of emitters in terahertz range; making conclusions about coverage of terahertz gap according to the maximum emission power of emitters recently developed; writing of the article.

- **V.** Analysis of possible improvements to the models of QCL simulation; calculations of triple-quantum-well transition times and current density versus applied electric field; review of the article.
Introduction

Semiconductor heterostructures have become a promising basis to produce emission in the higher end of terahertz region (1...10 THz). This thesis, divided into 4 chapters, handles modelling methods and algorithms for semiconductor quantum well heterostructures designed to produce radiation in the terahertz and mid-infrared (5...30 µm) region of the electromagnetic spectrum. The devices based on quantum well heterostructures are mainly produced using Molecular Beam Epitaxy (MBE), which is relatively expensive technology for small production quantities. Modelling of such devices is therefore an important stage in the development process to avoid faulty designs and save resources.

The vast increase in computing power during the last 10–15 years has provided opportunities to build more complex models to simulate the structures under conditions closer to the real environment. The models are based on the laws of quantum mechanics, covering many general phenomena (carrier scattering, phonon generation, electron transport, etc), being applied on quantum well heterostructures. Chapter 1 gives an overview of the theoretical principles of semiconductor emitters, covering also the main applications of terahertz range technology. Publication IV gives an overview of the latest achievements in this field as well.

The author’s main objective is to analyse, develop and optimise the selected models and algorithms used for quantum well heterostructures. The speed and accuracy of existing models should be improved. A cheaper way to produce incoherent terahertz radiation is analysed; therefore, a relevant new model for electron transport, where the electrons are pumped laterally, will be introduced. Lateral transport will give an incoherent terahertz radiation source that could work at room temperature, which makes it less expensive, as it does not require cooling. It is well known, that the majority of the sources otherwise require low working temperatures (below 164 K @ 3.0 THz in 2007 [1], 200 K @ 2.85 THz in 2012 [2]).

Chapter 2 presents the finding of a digitally graded parabolic quantum well heterostructure, which is necessary for the lateral electron transport task. Designing of parabolic quantum wells using MBE is an engineering challenge, as MBE grows structures layer by layer. In this chapter alterna-
tion of layers of 0.2825 nm GaAs and AlGaAs is used to achieve one of the subobjectives, the effective parabolicity. An analysis of numerical methods to solve the 1D time-independent Schrödinger equation is also performed in this chapter.

Chapter 3 covers the methodology for calculating the emission characteristics for emitters based on laterally pumped quantum well heterostructures. Generally the direction of electron transport is perpendicular to quantum wells, whereas in this chapter lateral transport is analysed. Relevant models for this are designed.

The models handled in this thesis are based on the Third Edition of "Quantum Wells, Wires and Dots" [3], which is also the main theoretical basis of this thesis. A set of computer codes [4] was published with the book, later referred to as QWWAD software tools, which is the realisation of the models described in the QWWAD book. Technically, it is a library containing general functions that are called by custom applications over a specific application interface (API). The set of codes is in continuous development by Leeds University and their partners, also being updated by the author of the thesis' new and modified models. References to some specific program in the QWWAD software tool are marked in bold letters.

However, some of the existing models are too time consuming. Numerical calculations carried out by the author in 2007 have shown long calculation times using either a personal computer (days) or even computing grid (hours), demanding improvements of algorithms and methods. Hence, another goal for the author is to redevelop some of the existing models in order to improve them by speed and accuracy. Optimisation without losing the precision of the results is a complex task. Chapter 4 gives an overview of improvements made to optimise the Quantum Cascade Laser simulation models (QCLSIM) in the QWWAD software tool.
List of acronyms

- AlGaAs - Aluminium Gallium Arsenide
- ARPACK - ARnoldi PACKage (numerical software library)
- EVP - Eigenvalue Problem
- EWC - Energy and Wave function Coupled solution method
- GaAs - Gallium Arsenide
- LAPACK - Linear Algebra PACKage
- LED - Light Emitting Diode
- MBE - Molecular Beam Epitaxy
- QCL - Quantum Cascade Laser
- QCLSIM - QCL Simulator (part of the software tools published as an appendix to QWWAD)
- QWWAD - Quantum Wells Wires and Dots (the book [3])
- RDX - Research Department Explosive (a type of explosive)
- RMS - Root Mean Square
Chapter 1

Overview of the terahertz region

This chapter is based on publication IV, providing an overview of advances in the terahertz range of the electromagnetic spectrum.

1.1 Overview of technology in the THz region

The range of the terahertz spectrum lies roughly between the frequencies of 1 THz and 10 THz (300 µm and 30 µm) [5], being between regions of mid-infrared and radio waves. The range of the terahertz region varies in different publications [6], giving the wider boundaries of the region 300 GHz and 30 THz, as seen in Fig. 1.1.

During the last 20 years, the development of sources and detectors of terahertz radiation has been rapidly boosted, providing numerous prospective applications of the terahertz range in spectroscopy, imaging, and communications [6]. The terahertz region of the electromagnetic spectrum has been underdeveloped since there was lack of technology generating coherent high-power radiation [7] (see also Fig. 1.2). In browsing one of the most commonly used scientific databases, the ISI Web of Knowledge, the number of publications in this region has increased significantly over the past two decades.

![Figure 1.1: The location of terahertz region in electromagnetic spectrum [6]](image-url)
of publications related to the terahertz area started growing from the beginning of the 1990s, making another vast "climb" between 2003 and 2009 (see Fig.1.3). The notable decrease in the number of publications after 2009 is presumably related to the global economic crisis.

The development of terahertz radiation sources has also been obstructed for a long time, due to low computing power available before the 1990s to do the modelling of quantum well heterostructures. This technology is said to be the most efficient technology to produce [6] and detect [11] coherent terahertz range emission. Vast development in computing technologies during the last decades have given way to develop applications based on quantum mechanical laws. The main necessity to have extreme computing power is to solve the Schrödinger equation, a differential equation that helps us to analyse the quantum well heterostructures as devices [3]. A variety of quantum well photodetectors of terahertz radiation is also being developed[12].

The majority of applications for terahertz radiation are related to spectroscopy, as all chemical elements have a unique absorption spectra at the terahertz region, which can be measured via absorption spectrum [13]. Security services need rapid detection and identification of all kind of explosives, illegal drugs, weapons, etc., which are excellent areas in which to apply
THz spectroscopy [14]. In addition, monitoring pollution and detection of noxious substances are some examples of such applications [15]. In Fig. 1.4, an example of spectra of absorption and transmittance of different explosives are shown.

Imaging is a second important application of the terahertz region. There are number of applications in medicine and security services to provide "seeing" through non-organic material (see Fig. 1.5) [15]. Terahertz pulsed imaging is a completely non-invasive and non-destructive way to analyse tablets in a pharmacy, performing 3D analysis on tablets to determine the coating integrity and thickness, and to detect and identify localised chemical or physical structure, etc. [15]. Non-destructive imaging of illegal drugs is also a good example of an application of terahertz imaging [16].

Terahertz communication is an area which promises very high bitrates in extremely absorption-free environments. Up to 1 Tbit/s is estimated by 2015 with wired communication, and 0.1 Tbit/s by 2020 with wireless [17].

1.2 THz radiation sources

The majority of sources of coherent terahertz radiation need to work at very low temperatures, which makes them expensive to use. Fortunately, not all of the applications need the coherent radiation, and can work with incoherent emission too. To get incoherent radiation, there is no need to use expensive lasers – some kind of cheaper solution can be developed that could
Figure 1.4: (a) Theoretical absorption spectrum of six different well known explosives. (b) Trasmittance of RDX explosive studied using four different Terahertz technologies from different sources. [14]

Figure 1.5: On the left side, there is a legendary photo of man with a knife, showing potential usage of terahertz rays [15]. On the right side, there is an example photo of analysis of a maple seed pod using THz-rays [5].
work even at room temperature. Incoherent radiation can be produced with a quantum well that is pumped by lateral current. The radiation is then generated in spontaneous radiative transitions between size-quantized states, requiring only the electron excitation to higher subbands. No population inversion is needed.

Quantum wells also have a broad range of applications. There have been AlGaAs/GaAs-based LEDs manufactured, using single and double heterostructures [18]. These heterostructures are used in highly effective red LEDs. A drawback of AlGaAs/GaAs-based LEDs is the requirement of very thin GaAs quantum wells surrounded by AlGaAs barriers. An example usage of quantum wells in LEDs is shown in Fig. 1.6.

![Figure 1.6](image-url) An example application of quantum wells in Light Emitting Diodes. The multiple quantum wells act there as an active region and increase the efficiency of light generation. Source: Fig. 4.14 (a) in [18]

The emission generated by a rectangular quantum well will have a relatively wide spectrum, being quite inefficient. To improve the bandwidth and make it comparatively narrow, a parabolic well should be used. Parabolic quantum wells have equal energy spacings and in such wells the only strong optical transitions are between adjacent states. The bandwidth of such emitters is limited by the spontaneous emission width.

In reality, it is a complex task to produce parabolic quantum wells, as the doping concentration needs to be changed homogeneously. One solution for producing parabolic wells could be by digital grading. This means that the shape of the parabolic well is replaced by rectangular layers of two materials giving the behaviour of a real parabolic quantum well (i.e. equidistant energy levels). Preparing such a digitised parabolic well is presented here in the second part of Ch. 2.

To make such a parabolic well emitting spontaneous radiation, external bias should be applied to it. External bias will produce a lateral electric field that gives additional kinetic energy to electrons. Additional kinetic energy increases the probability that electrons "jump" to higher energy levels via scattering process. This is therefore the excitation process via
lateral electric field and it is described in the second part of the thesis, in Ch. 3.

Using molecular beam epitaxy (MBE), a more complex and powerful radiation source can be built – a Quantum Cascade Laser (QCL) [9]. In 2002, the THz region was achieved using this technology [10]. In addition to MBE, Liquid Phase Epitaxy and Metal Organic Chemical Vapour Deposition methods are used to grow quantum well heterostructures with ultra-thin layers [19].
Chapter 2

Time-independent Schrödinger equation numerical solution methods. Application to digitally graded GaAs/AlGaAs parabolic quantum wells

This chapter is based on publications I and II, covering methodology to find electron energy eigenvalues and wave functions applied on digitally graded GaAs/AlGaAs parabolic quantum wells.

2.1 Introduction

To find the band structure of quantum well heterostructures, the basic time-independent 1D Schrödinger equation (see eq. (2.101) in [3]) is used:

$$-\frac{\hbar^2}{2m^*(z)} \frac{\partial^2}{\partial z^2} \psi(z) + V(z) \psi(z) = E \psi(z),$$  \hspace{1cm} (2.1)

where $m^*(z)$ is the effective mass that depends on the coordinate (whether it is GaAs or AlGaAs at that point), $\psi(z)$ is the wave function, and $E$ the energy eigenvalue. A more precise form of this equation with differential mass is preferable for thin potential barriers (see eq. (2.96) in [3]):

$$-\frac{\hbar^2}{2} \frac{\partial}{\partial z} \frac{1}{m^*(z)} \frac{\partial}{\partial z} \psi(z) + V(z) \psi(z) = E \psi(z). \hspace{1cm} (2.2)$$

The Schrödinger equation solutions are the energy eigenvalues and corresponding wave functions that allow us to make the electron transport calculations. Knowing energy eigenvalues helps us to find the sufficient design...
suitable for a digitally graded parabolic well that could give the best approach to the ideal parabolic well. It is known that parabolic wells give equally spaced energy levels; therefore, similar equal spacing must be found by constructing suitable digital grading. The solutions are also needed to perform the carrier transport calculations, which are described in the next chapter.

Two different methods to solve the Schrödinger equation were examined to find better accuracy. The first method attempted was the shooting method, and the other was the matrix eigenvalue and eigenstates solution method, both described in Paul Harrison’s book, "Quantum Wires, Wells, and Dots" [3].

2.2 Choosing a suitable numerical method

2.2.1 Shooting method

The shooting method is based on solving a differential representation of the Schrödinger equation with initial values that are known (see eq. (1.107) from [20]):

\[
\psi(z + \delta z) = \left[ \frac{2m^*(z)}{\hbar^2} (\delta z)^2 (V(z) - E) + 2 \right] \psi(z) - \psi(z - \delta z). \tag{2.3}
\]

In this differential equation it is clearly visible that three consecutive points in the wave function \( \psi \) depend on each other. Energy \( E \) is the parameter that is searchable and its value is being searched by boundary condition \( \psi(z \to \infty) \to 0 \). As the differential equation unites three points, two of them need to be given as the initial values. When these initial values are known, the final value for \( \psi(z_n) \) can be calculated by repeating the function iteratively \( 9n - 2 \) times.

Two initial values depend on the symmetry of the concrete wave function. The potential shape needs to be symmetrical to solve the Schrödinger equation using the shooting method. Solution of the differential equation needs to be started from the centre of potential. If the potential is symmetrical, the wave functions can either be symmetrical or non-symmetrical (see Fig. 2.1).

The initial values need therefore to be chosen according to the type of symmetry of the wave function. For symmetric and non-symmetric wave function, the initial values are (respectively):

\[
\begin{cases}
\psi(0) = 1 \\
\psi(\delta z) = 1
\end{cases}, \quad \text{and} \quad \begin{cases}
\psi(0) = 0 \\
\psi(\delta z) = 1
\end{cases}. \tag{2.4}
\]
As the shooting method uses a differential equation, the wave function is very precise near the initial values at $z = 0$ and $z = \partial z$. The final part of the wave function tends to deflect from the boundary condition $\psi(z \to \infty) \to 0$, as finding the precise $E$ that matches the real eigenvalue is very complicated. The real values always have a definite precision, which makes it impossible to find the precise value for energy. The fact that the final part of the wave function depends on the values in the beginning part of the wave function also makes sense. Therefore, it is very complicated to find solutions that are precise enough using the shooting method, as conformance to outer boundary conditions is not easily achievable.

Trials to solve the task using the shooting method are considered therefore to be too inaccurate. The orthogonality of wave functions were checked ($\int_{i \neq j} \psi_i \psi_j dz = 0$) and it was found that $\int_{i \neq j} \psi_i \psi_j dz \approx 0.1$ while $\int \psi_i \psi_i dz = 1$. That means the orthogonality of the wave functions is not good enough and this method needed to be replaced.

### 2.2.2 Energy and wave function coupled solution method (Newton method)

One way to improve the results of the shooting method was by applying another method to it. A method called the energy and wave function coupled solution method (EWC) was introduced [21] which is very precise but needs initial energies and wave functions as inputs. Those input values can therefore be prepared with the shooting method.

EWC method solves Schrödinger equations system with clearly fixed boundary conditions – both ends of the wave function can be fixed to some initial value. Both energy eigenvalues and wave functions are calculated simultaneously. The method is based on a 3-point scheme of spatial discretisation that corresponds to the equation derived from Schrödinger eq.
where $\psi_i$ denotes $\psi(z_i)$ and $V_i \equiv V(z_i)$. The exact boundary conditions for the outer calculation area of $\psi$ is set to zero (i.e. $\psi_1 = \psi_N = 0$). This boundary condition corresponds to the assumption of infinitive external barriers, as they cause the wave function to go to zero on the borders.

This is the representation with constant or slowly changing mass. In case of dynamic mass, it goes inside the brackets, as it can be seen in the following form:

$$
-\frac{\hbar^2}{2m} \left( \frac{\psi_{i+1} - \psi_i}{\Delta z} - \frac{\psi_i - \psi_{i-1}}{\Delta z} \right) \frac{1}{\Delta z} + V_i \psi_i = E \psi_i,
$$

(2.6)

where $m_{i+\frac{1}{2}}$ (and $m_{i-\frac{1}{2}}$) is the average mass between $m_i$ and $m_{i+1}$ ($m_i$ and $m_{i-1}$ respectively).

EWC method does not calculate the energies and wave functions directly, but only their corrections. That makes it more optimal, as it is not necessary to carry the absolute values through the calculations. The cycle of calculations is iterative, and can be stopped when the results no longer change much.

The eq. (2.6) can be converted to a representation that allows it to be calculated using the following equations:

$$
Y = \tilde{Y} + \delta Y, \quad (2.7)
$$

$$
\left[ \frac{\partial F}{\partial Y} \right] \times \delta Y = -\tilde{F}, \quad (2.8)
$$

where $\tilde{Y}$ is the approximate unknown vector, $\delta Y$ is the correction vector, $\tilde{F} \equiv (\tilde{F}_1, \tilde{F}_2, \ldots, \tilde{F}_N)^T$ is the right-hand side vector of the system calculated by $\tilde{Y}$, and $[\partial F/\partial Y]$ is the $N \times N$ Jacobi matrix with the Newton method derivatives [21]. The Jacobi matrix $[\partial F/\partial Y]$ has a triagonal structure where the first row and column are additionally filled out [21]. The first element of the main diagonal is zero. For example, for the constant mass formulation
shown in eq. (2.5), the matrix has the following structure:

\[
\left[ \frac{\partial F}{\partial Y} \right] =
\begin{bmatrix}
0 & 2\psi_2 & 2\psi_3 & 2\psi_4 & \cdots & 2\psi_{N-3} & 2\psi_{N-2} & 2\psi_{N-1} & 2\psi_N \\
\psi_2 & a_2 & c & 0 & \cdots & 0 & 0 & 0 & 0 \\
\psi_3 & c & a_3 & c & \cdots & 0 & 0 & 0 & 0 \\
\psi_4 & 0 & c & a_4 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\psi_{N-3} & 0 & 0 & 0 & \cdots & a_{N-3} & c & 0 & 0 \\
\psi_{N-2} & 0 & 0 & 0 & \cdots & c & a_{N-2} & c & 0 \\
\psi_{N-1} & 0 & 0 & 0 & \cdots & 0 & c & a_{N-1} & c \\
\psi_N & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1
\end{bmatrix},
\]

(2.9)

where \( c \equiv \hbar/(2m\Delta z^2) \) and \( a_i \equiv E - V_i - 2c \). For a deeper overview of this method, see [21].

This method is very precise, as by the end of iterations the increment vector \( \delta Y \) approaches computer zero. The disadvantages of this method are, firstly, the need to obtain an initial guess for the wavefunction, and secondly, the possibility to skip some eigenvalues. It was therefore decided not to use the method for rather difficult multibarrier digitised quantum well. The use of more common methods for trivial eigenvalue problems could solve these disadvantages, which is the topic of the next subsection.

**2.2.3 Method based on matrix eigenvalue standard solvers**

This method is based on classical linear algebra methods that are meant for finding eigenvalues and eigenvectors of matrices. This will be much more precise than the shooting method, because the elements no longer depend on each other. To solve the Schrödinger eq. (2.1) using matrices, the differential equation needs to be modified to difference representation (see eq. (3.53), page 83 in [3]):

\[
\frac{1}{m \ast (z + \delta z/2)} \psi (z + \delta z) = \\
= \left( \frac{2(\delta z)^2}{\hbar^2} [V(z) - E] + \frac{1}{m \ast (z + \delta z/2)} + \frac{1}{m \ast (z - \delta z/2)} \right) \psi (z) - \\
- \frac{1}{m \ast (z - \delta z/2)} \psi (z - \delta z).
\]

(2.10)

The eq. (2.10) is a difference equation, which means that it connects three consecutive wave function values. To solve it, standard eigenvalue and eigenvector computing methods can be used and the equation needs to
get matrix representation. The first step should be to rewrite the equation to the following form:

\[- \frac{1}{m^*} \left( z_{i-\frac{1}{2}} \right) \psi_{i-1} + \left( \frac{2(\delta z)^2}{\hbar^2} V(z_i) + \frac{1}{m^* \left( z_{i+\frac{1}{2}} \right)} + \frac{1}{m^* \left( z_{i-\frac{1}{2}} \right)} \right) \psi_i - \]

\[- \frac{1}{m^* \left( z_{i+\frac{1}{2}} \right)} \psi_{i+1} = \frac{2(\delta z)^2}{\hbar^2} E \cdot \psi_i , \quad (2.11)\]

where \( \psi_i = \psi(z) \), \( \psi_{i+1} = \psi(z \mp \delta z) \), \( z_i = z \) and \( z_{i \mp \frac{1}{2}} = \frac{z \pm \delta z}{2} \). For eigenvalue problems it is better to present this equation without the coefficient in front of energy \( E \). Therefore, the whole equation needs to be multiplied by \( \frac{\hbar^2}{2(\delta z)^2} \) and will get the following form:

\[- \frac{k}{m^* \left( z_{i-\frac{1}{2}} \right)} \psi_{i-1} + \left( \frac{k}{m^* \left( z_{i+\frac{1}{2}} \right)} + \frac{k}{m^* \left( z_{i-\frac{1}{2}} \right)} \right) \psi_i - \]

\[- \frac{k}{m^* \left( z_{i+\frac{1}{2}} \right)} \psi_{i+1} = E \cdot \psi_i , \quad (2.12)\]

where \( k \) denotes the coefficient \( \frac{\hbar^2}{2(\delta z)^2} \).

The coefficients before \( \psi \)-s can be placed into a symmetric tridiagonal band matrix and the task can be reformulated as a standard eigenvalue problem:

\[[A] \cdot [\psi] = E [\psi] , \quad (2.13)\]

where \([A]\) is the \( N \times N \) matrix, \([\psi]\) is a column vector with \( N \) elements, and \( E \) is the energy eigenvalue we want. By solving this eigenvalue problem with standard software (e.g. \textit{dstev} and \textit{dstevx} in LAPACK\(^1\)), the energies \( E_l \) (eigenvalues) and corresponding wave functions \( \psi_l \) (eigenvectors) can be found (where \( l \) is the number of energy level, \( l = 1 \ldots N_l \)).

The orthogonality of the wave functions is much better using this method. The value of nondiagonal elements of the orthogonality matrix was \( \int_{i \neq j} \psi_i \psi_j dz \approx 10^{-9} \), which shows clearly how much better this method is than the shooting method.

\(^1\)LAPACK is an acronym of words \textit{Linear Algebra PACKage}. See http://www.netlib.org/lapack for more information.
### 2.2.4 Comparison of all three methods

The comparison of the main properties and pros and cons of all these methods are listed in tables 2.1 and 2.2.

#### Table 2.1: Comparison table of all three methods. $N$ is the number of net points

<table>
<thead>
<tr>
<th>Property</th>
<th>Shooting method</th>
<th>Newton method</th>
<th>Matrix eigenvalue method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Boundary conditions</td>
<td>One side is set to 2 pts</td>
<td>Both sides fixed to 0</td>
<td>Neither side fixed</td>
</tr>
<tr>
<td>2. Wave function orthogonality</td>
<td>$\approx 10^{-1}$</td>
<td>$\approx 10^{-12}$</td>
<td>$\approx 10^{-9}$</td>
</tr>
<tr>
<td>3. Time efficiency</td>
<td>Medium (time $\sim N$)</td>
<td>Very High (time $\sim N$)</td>
<td>High (time $\sim N^3$)</td>
</tr>
<tr>
<td>4. Initial guess for wave function</td>
<td>Unnecessary</td>
<td>Necessary</td>
<td>Necessary</td>
</tr>
<tr>
<td>5. Implementation complexity</td>
<td>Medium</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>6. Traceability</td>
<td>Good</td>
<td>Poor</td>
<td>Poor</td>
</tr>
<tr>
<td>7. Other assumptions</td>
<td>Symmetrical potential needed</td>
<td>Outer barriers are infinitively high</td>
<td>Not very clear boundary conditions</td>
</tr>
</tbody>
</table>

#### Table 2.2: Advantages and disadvantages of the three methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shooting method</td>
<td>Easy to implement</td>
<td>Not very precise; May skip energy levels if they are too close to each other</td>
</tr>
<tr>
<td>Newton method</td>
<td>Boundaries fixed to 0; Relatively fast</td>
<td>Initial guess for energies and wave functions needed; May skip energy levels if they are too close to each other</td>
</tr>
<tr>
<td>Matrix eigenvalue method</td>
<td>Standard implementation</td>
<td>Not very clear boundary conditions</td>
</tr>
</tbody>
</table>
Figure 2.2: An example of parabolic quantum well with digital grading. Minimal width of a monolayer is 0.2825 nm due to restrictions set by MBE [3].

2.3 Optimising digitally graded potentials

2.3.1 Introduction

The main task in this chapter is to find digitally graded potential layout that could be a good approximation to parabolic quantum well. An example of parabolic quantum well with digital grading is shown in Fig. 2.2. Digital grading is needed to simplify the production process of parabolic quantum well devices. Molecular Beam Epitaxy is a technology that layers semiconductor materials one on top of the other [22]. Using this technology, it is not possible to manufacture quantum wells that have a parabolic shape. One option to overcome this obstacle is digitalisation of the parabolic well. This means the two materials used in the parabolic well (GaAs and Al$_x$Ga$_{1-x}$As) are altered as many times as necessary to achieve similar behaviour to true the parabolic well. Finding the right digital grading can be time consuming, because all the thick layers must be carefully shifted left and right to find the best approach to equal spacing between energy levels. The spacing serves as a feedback in the digitised parabolic well construction method.

2.3.2 Building initial potential

Building the initial structure was relatively easy. The first criterion while building the structure was the integral of the potential function (the area) that needs to be equal for both cases – for the original parabolic well and for the digitised well. This defined the number of layers. The second criterion was needed to define their initial placement. The algorithm for assigning the layers to their initial places started moving from the centre of potential and integrated over the original potential well. If the integral happened to exceed the area of one monolayer, a layer was put into this place where it
occurred. Then the counter was zeroed and integration continued until the end of the potential. Initial potential behaved more or less like the parabolic well, but it was still not very precise.

### 2.3.3 Optimising the potential

To optimise the initial potential, several algorithms were tried. The first and easiest was to move layers one step left and right, doing it one by one with each of the layers. After each movement, the unified spacing parameter between energy levels was calculated again and compared to the current minimal spacing parameter. This process was also continued with several other algorithms until the spacing no longer improved.

To characterise the deviation of spacings from desired value, a statistical parameter – Root Mean Square (RMS) – was used. The parameter was calculated using the following equation:

\[
RMS = \sqrt{\frac{\sum_{i=1}^{N-1} (\Delta E_{i,i+1} - \Delta E_{\text{wanted}})^2}{N-1}},
\]  

(2.14)

where \(\Delta E_{i,i+1} = E_{i+1} - E_i\) is the spacing between energy levels \(i\) and \(i + 1\), \(\Delta E_{\text{wanted}}\) is the desired energy spacing (a constant value), and \(N\) is the number of energy levels. After each displacement of a layer, the \(RMS\) is recalculated and used in later movements as a comparison. If a new relocation changed the \(RMS\), the position of the corresponding monolayer is fixed and another layer taken into focus.

All the algorithms used are described below and shown in Fig. 2.3.

---

\(\text{the width of a monolayer (0.2825 nm)}\)
A. Moving one stick at a time to the left and right  This was the first method tried to improve the RMS of energy level spacings. The algorithm started from the centre of potential and moved symmetrically to the direction of the edge. When it reached the end, it started coming back in reverse to the centre. The algorithm continued this cycle until the movement of layers no longer improved the energy spacings.

If the new position of a layer is occupied by another layer, then this layer will also be moved on. If there are many layers, then all of them will be moved.

B. Moving a group of sticks to the left and right  This method was the first improvement to the previous one, as it did not give the best approach. The method starts again from the middle of the well and goes to the side. The first movement incorporates all the layers that are moved to the right and to the left. The second movement leaves out the first layer and takes all the others \((n - 1)\) if \(n\) is the number of layers) with it.

C. Moving two consecutive sticks towards and away from each other  This was the last improvement attempted to make the achieved results even better. The process starts again from the middle of the well and goes to the side. All the consecutive layers were moved towards each other and then away from each other.

In final calculations, all the methods were combined, alternating them after each step. Acting that way gave the best result – the precision of energy spacings came to around 5 percent, which was also the initial goal [23]. See Fig. 8 in [23].

2.4 Conclusion

There are two main results achieved in this chapter. The aim of the first task was to find a proper method for band structure calculation. This task was performed successfully – a standard linear algebra method to find matrix eigenvalues and eigenvectors was chosen to be the best one. The initial data was chosen such that the eigenvalues were the energy values and eigenvectors the corresponding wave function.

The second task in this chapter was to find a good approximation to parabolic quantum well using digital grading, i.e. alteration of two different substances, GaAs and AlGaAs. After the layout of layers was fixed, the energy levels had more or less equal spacing, and the final results were to be calculated. After the final energy eigenvalues and wave functions were ready, the process continued with carrier transport calculations, which is handled in the next chapter 3.
Chapter 3

Lateral transport task of a quantum well-based broadband terahertz emitter

This chapter is based on publication III, covering methodology to calculate emission rates for lateral transport of charge carriers.

3.1 Introduction

The quantum wells we are using are two dimensional, which means that the electrons have two directions where they can move and one direction where they are fixed. This is therefore called in-plane electron transport. Fig. 3.1 shows how the electrons can move along the valley (y-axis) and up along the subband (x-axis). Along the z-axis the electrons are not free.

Lateral transport brings up the effect of scattering – the electrons are colliding with vibrating atoms in the lattice (phonons) and other electrons. Such collisions heat the electrons up, giving them higher kinetic energy. This will cause the electrons to climb up along the subband until they collide with the lattice and jump to higher subbands through the scattering process (see Fig. 3.2). Higher subbands will therefore have a higher population of electrons. Electrons in excited subbands will then relax into lower ones and produce spontaneous emission of photons. The main idea is therefore in increasing the output power by additional spontaneous emission.

A device emitting terahertz radiation can be built using this idea. An example of such a device can be seen in Fig. 3.3.

The lateral electric field therefore causes the Fermi-Dirac distribution function needed in transport calculations to shift along the energy axis to some extent (see Fig. 3.4 and section 3.2). The original distribution function for equilibrium is given by (see eq. (2.49) in [3]):
Figure 3.1: In-plane dispersion curves and the subband structure. Source: Fig. 2.5, page 22 in [3].

Figure 3.2: Model of subband excitation and relaxation processes. Nonradiative intersubband scattering processes (polar LO phonons and acoustic deformation potential phonons) cause electron transitions between subbands. The lateral electric field accelerates electrons within every subband. Optical radiation output is caused by spontaneous drop of electrons from higher subbands to lower ones, while $\Delta k \approx 0$.

Figure 3.3: An example device based on a laterally pumped quantum well
\[ f^{FD}(E) = \frac{1}{\exp\left[\frac{(E - E_F)}{kT}\right] + 1}, \quad (3.1) \]

where \( E_F \) is the Fermi energy.

Figure 3.4: The shifted Fermi-Dirac distribution function with effective electron temperature. The shift is caused by the lateral electric field, and the flatness is caused by the increase in temperature.

### 3.2 Shifted Fermi-Dirac distribution function

The internal electric field of lateral transport causes electrons to obtain a remarkable drift velocity \((v_d > 0)\), heating themselves up. This will cause the Fermi-Dirac distribution function (3.1) to shift along the \( E \) axis by energy that corresponds to \( v_d \), as \( E_{shifted} = E_0 + \frac{\hbar^2 k^2}{2m^*} = E_0 + \frac{\hbar^2 v_d^2}{2m^*} \).

The distribution function will then get the following form:

\[
f^{sFD}(k) = \left[ 1 + \exp\left\{ \frac{E_{n0} + \frac{\hbar^2 ((k_x-k_{x0}(F,T_{latt}))^2+k_y^2)}{2m^*} - E_{F_n}}{k_B T_{el}(F,T_{latt})} \right\} \right]^{-1}, \quad (3.2)
\]

where \( k = (k_x,k_y) \) is the in-plane wave vector that is proportional to energy \( E \) in the original equation, \( E_{n0} \) is the subband minimum energy, \( E_{F_n} \) is the quasi-Fermi level of the \( n \)-th subband, \( k_{x0} \) is the drift wave vector along the \( x \)-axis, and \( T_{el} \) the heated electron temperature. \( k_{x0} \) and \( T_{el} \) depend both on the electric field \( F \) and lattice temperature \( T_{latt} \).

Two unknown parameters are introduced with *shifted* distribution function – drift wave vector and electron temperature \((k_{x0} \text{ and } T_{el})\) – and both are functions of applied electric field \( F \) and lattice temperature \( T_{latt} \). The author decided to take these functions from the literature, as there have already been several studies done on this matter.

Dependence between the drift velocity and electric field is taken from [24]. The relevant figure is 3.5 (the left one), which is a copy of the original figure from the paper. The points on this graph were carefully written out.
and interpolated using Lagrange’s interpolation method. The interpolated data were then written out again and presented here as the right curve in the same figure. It can be seen that the result conforms more or less to the original data. Experiments showed that this small difference does not affect the overall result significantly. In calculations where the wave vector $k$ is used instead of speed $v$, the conversion can be done using the following relationship [25]:

$$k(F) = \frac{m^*v(F)}{\hbar}.$$  (3.3)

The electron temperature and electric field dependence are also taken from the same paper. Fig. 4 in [24] shows the dependence between the electric field and average energy for GaAs at 300 K. The comparison is presented in Fig. 3.6, where both the original and interpolated curves are presented. The average energy $E$ is converted to temperature $T$ using the equation from the ideal gas model [25]:

$$T(F) = \frac{2E(F)}{3k_B}.$$  (3.4)

### 3.3 Calculation methodology

#### 3.3.1 Overview

The process of calculations is described in Fig. 3.7. The calculation has two cycles: the inner one is iterative to find the proper subband populations, and the outer one that is over the electric field values. The band structure needs to be calculated in advance.

#### 3.3.2 Calculation of raw scattering rates

Electrons moving in a crystal lattice will sooner or later collide with lattice atoms. That way the electrons can change their states – collisions may either increase or decrease their energy. According to Fermi’s Golden Rule, the scattering process is described as follows: if there is a time-dependent perturbation $\hat{\mathcal{H}}$, which could transfer an electron from state $|i\rangle$ with energy $E_i$ to state $|f\rangle$ with energy $E_f$, the lifetime of the carrier in state $|i\rangle$ is (according to eq. (10.1) in [3]):

$$\frac{1}{\tau_i} = \frac{2\pi}{\hbar} \sum_f \left| \langle f | \hat{\mathcal{H}} | i \rangle \right|^2 \delta (E_f - E_i).$$  (3.5)

There are several types of scatterings, but only longitudinal optic and acoustic deformation potential are taken into account, as the others do not affect the results greatly. All the scattering types have their own equations
Figure 3.5: Comparison of the original (left) and interpolated (right) dependence between drift velocity and electric field at 300 K in GaAs (source: Fig. 3 in [24]).

Figure 3.6: Comparison of the original and the interpolated dependence between average energy and electric field at 300 K in GaAs (source: Fig. 4 in [24]).
Figure 3.7: Overview of the calculation process. The source of codes' given mean: *original* – code taken from [3], *modified* – code taken from [3] and modified by the author, and *new* – code written by the author. Variable $N$ means the total number of energy levels (subbands). The majority of the calculations are automated using a program called *multicalculate*.
for the perturbation $\tilde{H}$, which gives the final form to the scattering rate equations described below.

The calculation of longitudinal optic (LO) phonon scatterings were performed using Paul Harrison’s book (section 10.3 in [3]). According to this section, the scattering rate eq. (3.5) can be improved and will finally get the following form (see eq. (10.147) in [3]):

$$\frac{1}{\tau_i} = \frac{m^* e^2 \omega}{(2\pi)^2 \hbar^2} \int_{-\infty}^{\infty} \frac{\pi |G_{if}(K_z)|^2}{\sqrt{K_z^2 + 2K_z^2 (2k_i^2 - \frac{2m^*\Delta}{\hbar^2}) + (\frac{2m^*\Delta}{\hbar^2})^2}} dK_z, \quad (3.6)$$

where $\Delta$ is the sum of subband minimal energy and the kinetic energy within the band $E_f - E_i \pm \hbar \omega$ (The upper sign in front of $\hbar \omega$ represents the emission of a phonon and decreases the absorption), $P = \left(\frac{1}{\epsilon_{\infty}} - \frac{1}{\epsilon_{s}}\right) \left(N_0 + \frac{1}{2} \mp \frac{1}{2}\right)$ (where $\epsilon_{\infty}$ and $\epsilon_{s}$ are the high- and low-frequency permittivities of the material, and $N_0 + \frac{1}{2} \mp \frac{1}{2}$ represents the number of phonons per unit area within the crystal, having minus in case of emission and plus in case of absorption, where $N_0$ is the Bose-Einstein factor), $K_z$ and $\omega$ are the wave vector (along the growth axis) and angular frequency of the phonons, $k_i$ is the momentum of phonon in the initial state, and $G_{if} = \int \psi_f^*(z) e^{-iK_z z} \psi_f(z) dz$ is the form factor of scattering events.

Scattering rate acoustic deformation potential (AC) calculations were performed according to section 10.9 in Paul Harrison’s book, [3]. The equation for AC scattering is in the form (see eq. (10.186) in [3]):

$$\frac{1}{\tau_i}(k_i) = \frac{D_A^2 m^*}{\rho v_s (2\pi)^2 \hbar^2} \left(N_0 + \frac{1}{2} \pm \frac{1}{2}\right) \int_0^\infty \int_0^{2\pi} (G_{if}(K_z))^2 \times$$

$$\times \left( \Theta(\alpha_1) \frac{\alpha_1 \sqrt{\alpha_1^2 + K_z^2} + \Theta(\alpha_2) \frac{\alpha_2 \sqrt{\alpha_2^2 + K_z^2}}{\alpha_1 - \alpha_2}} \right) d\phi dK_z, \quad (3.7)$$

where $D_A$ is electron acoustic deformation potential (in case of $\Gamma$-valley of GaAs $D_A = 7.0$ eV. Source: Table 2.1 in [26]), $G_{if}$ is the form factor of scattering events (see the previous paragraph), $\alpha_{1,2} = -k_i \cos \phi \pm \sqrt{k_i^2 \cos^2 \phi - \frac{2m^*\Delta E}{\hbar^2}}$ (according to eq. (10.182) in [3]), $k_i$ is the phonon wave number of initial subband, $K_z$ is the wave vector of the phonons, $\Theta$ is the Heaviside function, $\rho$ and $v_s$ are the density and speed of sound (respectively) in GaAs, and $N_0 + \frac{1}{2} \mp \frac{1}{2}$ represents the number of phonons per unit area within the crystal, having minus in case of emission and plus in case of absorption (where $N_0$ is the Bose-Einstein factor).

There is a standard implementation to solve these tasks [4]. A program called srelo is designed for LO scattering calculations and sradp for AC scattering. Therefore, in the overall calculation process, this part was solved using these standard programs.
3.3.3 Calculation of quasi-Fermi energies

In the calculation of mean scattering rate (subsection 3.3.4), the shifted Fermi-Dirac distribution function is constructed. According to the eq. (3.1), the Fermi-Dirac distribution function depends on Fermi energy $E_F$. In lateral transport the Fermi energy is separate for all the subbands:

$$f_{FD}^i(E) = \frac{1}{\exp[(E - E_{Fi})/kT] + 1},$$  

(3.8)

where the index $i$ expresses the index of subband. Fermi energy is therefore a ‘quasi’ energy describing the carrier population within a subband.

Quasi-Fermi energies need to be calculated for the shifted distribution function in the next subsection. As the Fermi energy is tightly related to carrier population within one subband, the problem can be solved using the eq. (2.48) in [3]. The equation will give the electron occupation of state $i$:

$$n_i = \frac{m^*}{\pi \hbar^2} \int_{\text{subband}} f_{FD}^i(E) \, dE.$$  

(3.9)

By putting equations (3.8) and (3.9) together, the Fermi energies can be found using reverse search method if the subband populations are known. The subband populations will be calculated later in subsection 3.3.5. This is not a problem that they are calculated in reverse order, because of the iterative calculation procedure. Arbitrary populations are used in the first round.

In Paul Harrison’s book, [3], the method was implemented for equilibrium (a program called sbp). The script was modified to provide a shifted Fermi-Dirac distribution function for the non-equilibrium.

3.3.4 Calculation of mean scattering rates

Calculation of subband populations in the next subsection (3.3.5) depends on mean scattering rates. Therefore, it is necessary to find the mean rates using eq. (8.149)$^1$ from Paul Harrison’s book, [3]:

$$\langle \frac{1}{\tau_{if}} \rangle = \frac{\int \frac{1}{\tau_{if}} f_{FD}^i(E) \left(1 - f_{FD}^f(E + E_{\text{phonon}})\right) \, dE}{\int f_{FD}^i(E) \, dE},$$  

(3.10)

where the indexes $i$ and $f$ stand for ‘initial’ and ‘final’ states, and $E_{\text{phonon}}$ is the phonon energy with minus in case of emission and plus in case of absorption.

As lateral transport shifts the distribution function away from the zero, its shifted form should be used in this equation too. The integral over energy

$^1$The corrected form is presented here according to the errata of the book
has also to be changed to integral over wave vector $k$ then. In computational implementation, the equation will look like this:

$$
\left\langle \tau_{if}^{-1} \right\rangle = \frac{\int_{k_x} \int_{k_y} \frac{1}{\tau_{if}} f^{sFD}_i(k_x,k_y) \left( 1 - f^{FD}_f(E - \mp E_{phonon}) \right) dk_x dk_y}{\int_{k_x} \int_{k_y} f^{sFD}_i(k_x,k_y) dk_x dk_y},
$$

(3.11)

where $E = E_i + \frac{\hbar^2}{2m^*} \left( k_x^2 + k_y^2 \right)$ (where $E_i$ is the subband minimum energy and $\frac{\hbar^2}{2m^*} \left( k_x^2 + k_y^2 \right)$ the kinetic energy), and $f^{sFD}_i(k)$ is the shifted Fermi-Dirac distribution function (see section 3.2).

This function was implemented using a standard program **srmpr** from Paul Harrison’s book, [3], which was improved with the shifted distribution function.

### 3.3.5 Calculation of subband populations using rate equations

To calculate the subband populations, rate equations need to be constructed using average scattering rates. It is known that the number of electrons leaving a state is equal to the number of electrons coming to the state. Therefore, the following rate equation can be used:

$$
\frac{dn_f}{dt} = \sum_{i=1}^{N} \frac{1}{\tau_{if}} n_i - n_f \sum_{i=1}^{N} \frac{1}{\tau_{fi}} = 0,
$$

(3.12)

where $\tau_{if}^{-1}$ is the total averaged scattering rate from $i$-th to $f$-th subband, $n_i$ is the electron population of $i$-th subband, and $N$ the number of subbands. This equation contains $N$ unknown variables $n_i$-s, which means at least $N - 1$ equations need to be found in addition to solve it. As the equation is meant for one subband only ($f$-th), it can be applied for all the $N$ subbands:

$$
\begin{align*}
\sum_{i=1}^{N} \frac{1}{\tau_{i1}} n_i - n_1 \sum_{i=1}^{N} \frac{1}{\tau_{i1}} &= 0 \\
\sum_{i=1}^{N} \frac{1}{\tau_{i2}} n_i - n_2 \sum_{i=1}^{N} \frac{1}{\tau_{i2}} &= 0 \\
&\vdots \\
\sum_{i=1}^{N} \frac{1}{\tau_{iN}} n_i - n_N \sum_{i=1}^{N} \frac{1}{\tau_{iN}} &= 0
\end{align*}
$$

(3.13)

Unfortunately, in this case, zero populations will also give a true result ($n_i = 0$, $i = 1 \ldots N$). Therefore, an additional equation which connects the subband populations to overall electron concentration in the semiconductor should be introduced:

$$
n_1 + n_2 + n_3 + \cdots + n_N = N_D,
$$

(3.14)

where $N_D$ is the overall density of electrons.
These equations can easily be solved using the following matrix equation

\[ Ax = y, \quad (3.15) \]

where

\[
A = \begin{bmatrix}
\sum \frac{1}{\tau_{11}} & \frac{1}{\tau_{12}} & \cdots & \frac{1}{\tau_{1(N-1)}} & \frac{1}{\tau_{1N}} \\
\frac{1}{\tau_{21}} & \sum \frac{1}{\tau_{22}} & \cdots & \frac{1}{\tau_{2(N-1)}} & \frac{1}{\tau_{2N}} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{1}{\tau_{(N-1)1}} & \frac{1}{\tau_{(N-1)2}} & \cdots & \sum \frac{1}{\tau_{(N-1)i}} & \frac{1}{\tau_{N(N-1)}} \\
\frac{1}{\tau_{1(N-1)}} & \frac{1}{\tau_{2(N-1)}} & \cdots & \frac{1}{\tau_{(N-1)(N-1)}} & \frac{1}{\tau_{N(N-1)}}
\end{bmatrix},
\]

\[
x = \begin{bmatrix}
n_1 \\
n_2 \\
\vdots \\
n_{N-1} \\
n_N
\end{bmatrix}, \quad \text{and} \quad y = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
N_D
\end{bmatrix}.
\]

To solve this task, a new program ND2ni was written.

A similar model for finding subband populations can also be used in quantum well infrared photodetectors [27].

### 3.3.6 Calculation of emission characteristics

The emission characteristic we are interested in is the emitted power, both the total value and the spectrum. The whole cycle of calculations done before was to find the subband populations \( n_i \) for the equation of emitted power:

\[ P_{\text{total}} = \sum_{i>f}^{N} \frac{n_i}{\tau_{i_f}} (E_i - E_f), \quad (3.16) \]

where \( n_i \) is the population of \( i \)-th subband, \( \tau_{i_f} \) is the spontaneous radiative lifetimes and \( h\omega_{i_f} = E_i - E_f \) the energy difference between \( i \)-th and \( f \)-th subbands [3]. The radiative lifetime is given by:

\[ \frac{1}{\tau_{i_f}^{\text{rad}}} = \frac{e^2\bar{n}}{3\pi\varepsilon_0 e^3 \hbar^4} \left( E_i - E_f \right)^3 d_{i_f}^2, \quad (3.17) \]

where \( \bar{n} \) is the refractive index, \( d_{i_f} \) are the optical dipole matrix elements, \( d_{i_f} = \int \psi_f(z) z\psi_i(z) \, dz \) [3]. Spontaneous radiative lifetimes \( \frac{1}{\tau_{i_f}^{\text{rad}}} \) do not depend on subband populations \( n_i \), which means they can be calculated prior to other calculations. This is important in practical implementation,
because this calculation can be omitted in the main cycle to save time. The second output needed, the power spectrum, is calculated as a well-known Lorentzian spectrum:

\[
P(E) = \sum_{i>f}^{N} \frac{n_i}{\tau_{i,f}^{rad}} (E_i - E_f) \frac{\Gamma}{\pi} \frac{1}{(E - (E_i - E_f))^2 + \Gamma^2},
\]

(3.18)

where \(\Gamma\) is the line width (half-width at half maximum) of the intersubband transitions, \(E_i\) and \(E_f\) the energies of initial and final subbands, and \(\tau_{i,f}^{rad}\) the spontaneous radiative lifetimes.

### 3.3.7 Overall calculation automation

According to Fig. 3.7, the main cycle of transport calculations are being controlled by the program `multicalculate`. This means that the band structure needs to be calculated in advance, then the transport calculations can be started. The program requests all the input data from the user (i.e. the electric field values \(F_i\), temperature \(T\), and total electron density \(N_D\)) and calls the subroutines automatically with the right arguments in the right order. The program checks the change in power emission after each inner iterative calculation cycle and exits as soon as one per cent of precision is achieved. All the inner iterative cycles are a part of an outer cycle which runs over the array of electric fields.

### 3.4 Conclusion

The calculations show that the emitted radiation spectrum peaks are above the blackbody level for all considered lattice temperatures 77–400K [28], [29]. The main problem through the development process was validation of the results. As there are standard implementations (that can be trusted) for an unbiased case in Paul Harrison’s book, [3], the results for \(F = 0\ V/cm\) can be compared. Calculations performed for the 7 THz frequency show the difference under 1\%. The comparison is given in table 3.1. During the early stage of calculations, the error in results was around 30\%. By now, the methods have been improved and corrected, so the precision has become better.

It can therefore be stated that the modelling method for electron transport in laterally pumped semiconductors has been developed and implemented.
Table 3.1: Comparison between results got by standard implementation (taken from [3]) and author’s implementation. Electric field is not applied \((F = 0 \text{ V/cm})\), having only black body radiation.

<table>
<thead>
<tr>
<th></th>
<th>77 K</th>
<th>300 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation taken from QWWAD ([3]). Averaging is done using eq. (3.10)</td>
<td>16.97 W/m²</td>
<td>16.98 W/m²</td>
</tr>
<tr>
<td>Implementation made by the author. Averaging is done using eq. (3.11)</td>
<td>17.02 W/m²</td>
<td>17.07 W/m²</td>
</tr>
<tr>
<td>Deviation</td>
<td>0.3 %</td>
<td>0.5 %</td>
</tr>
</tbody>
</table>
Chapter 4

Optimisation of mathematical models for terahertz range quantum cascade lasers

This chapter is based on publication V and manuscript [30], analysing strategies for optimisation of mathematical models.

4.1 Introduction

In this chapter, possibilities for optimisation of mathematical models in the QWAAD software tool [4] are analysed, particularly models in the QCL simulator (QCLSIM), which is a module of this tool. The software tool is developed by Leeds University and published as an appendix to Paul Harrison’s book, "Quantum Wells, Wires and Dots" (QWAAD). It is designed to simulate semiconductor nanostructures. The software has a modular architecture, providing easy understanding of the overall calculation process, while each module may have a very complex structure. Although the toolkit is a promising set of software programs, there is a growing need for improving the speed and accuracy of components of the toolkit. For large quantum well structures, the toolkit is rather used on computing grids, as the computational time is overwhelmingly large. Users need to run this software on a regular PC to perform the simulations fast enough (in the range of 10 minutes, rather than hours). In this chapter, the options regarding how to improve the toolkit are analysed and software updates are implemented. The main program in the QCLSIM toolkit is called qclsim, which is described in Fig. 4.1. Several sub-programs shown in that scheme contain other sub-programs. For example, the program called poisson Schroedinger contains scripts visible in Fig. 4.2 and the program called solve Schroedinger has content visible in Fig. 4.3.
Calculate initial estimate of wave functions using flat electron distribution

Specify the magnitude of electric field and lattice temperature

Generate alloy and doping profiles for 4−period structure

Generate conduction band potential, effective masses and nonparabolicity parameter

Find internal electric field (for GaN only)

Find total electric potential in which an electron is confined

Generate total potential for 4-period structure

Change 4−period structure into a single period structure with 1−period thick barriers (in case of dense matrix method)

Calculate initial estimate of wave functions using flat electron distribution

Use fixed $T_e$?

Yes

No

Specify the magnitude of electric field and lattice temperature

Generate alloy and doping profiles for 4−period structure

Generate conduction band potential, effective masses and nonparabolicity parameter

Find internal electric field (for GaN only)

Find total electric potential in which an electron is confined

Generate total potential for 4-period structure

Change 4−period structure into a single period structure with 1−period thick barriers (in case of dense matrix method)

Calculate initial estimate of wave functions using flat electron distribution

Use fixed $T_e$?

Yes

No

Estimate self−consistent electron temperature using Brent method

Solve the Poisson and Schroedinger equations self−consistently

Is $T_e$ converging enough?

Yes

No

Solve the Poisson and Schroedinger equations self−consistently

Generate plot files for wave functions

Use dense matrix method?

Yes

No

Generate gain data

Generate current data

Generate gain and current data using dense matrix method

End

End

Figure 4.1: Main architecture of Quantum Cascade Laser toolkit. The bold strings inside the large dotted rectangle are all sub-programs inside the main toolkit program qclsim.
Figure 4.2: Architecture of Poisson-Schrödinger equation solver (poisson_schroedinger)

Figure 4.3: Architecture of Schrödinger solver (solve_schroedinger)
Enter input parameters
1. Electric field values
2. Temperature values
3. Material name
4. Number of periods
5. Other execution parameters
6. Data reuse parameters

Run QCLSIM toolkit
Are all electric field and temperature values passed?
End

Figure 4.4: Overall wrapper of qclsim

find_scattering_rates
Initialisation
Generate array of required rates
If material = SiGe
find_intervalley_phonon_rates
Generate intervalley phonon s.r.
Generate acoustic phonon s.r.
Generate alloy disorder scattering rate
Generate interface roughness s.r.
Generate impurity scattering rate
End

Other materials

Figure 4.5: Architecture of scattering rate calculation module (find_scattering_rates)
During the optimisation task, two structures are used to perform the test calculations to verify the results. The first structure is a single quantum well (10 nm GaAs/Al\textsubscript{0.15}Ga\textsubscript{0.85}As quantum well with 50 nm barriers), later referred to as structure #1. The second structure is a triple-quantum-well GaAs/Al\textsubscript{0.2}Ga\textsubscript{0.8}As QCL described in [3] (pages 337–338), labelled as structure #2, which can be seen in Fig. 4.6. In tables 4.1 and 4.2 you can see the time spent on running the QCLSIM tool for each module. This detailed analysis helps us to discover the most inefficient modules in the software that could be optimised. There are two methods for solving the Schrödinger equation used (see module f\textbf{w}f in Fig. 4.3). Method #1 is the Taylor approximation of finite difference method, which is fast and accurate for states near the band edge, but gives errors when $E_{state} - E_{band\text{edge}} \approx E_{bandgap}$ [3]. Method #2 is the finite difference method accounting for band nonparabolicity, which has good precision for all states but is slow [31]. In this analysis, only structure #1 was calculated, as the overall process needs the device to be nothing other than QCL.

According to the results of experimental calculations\textsuperscript{1} given in table 4.1, the most time-consuming software tool is \texttt{self\_consistent\_elec\_temp}.

\textsuperscript{1} All the calculations in this chapter were done using PC with the processor: 64-bit Intel(R) Core(TM) i5-3470 CPU @ 3.20GHz.
Table 4.1: Analysis of time consumption per sub-programs in qclsim. The number of calculations was 3 for method #1 and 1 for method #2. It can be seen that the most time-consuming sub-program is self_consistent_elec_temp.

<table>
<thead>
<tr>
<th>Sub-program of qclsim</th>
<th>Schrödinger solution method #1</th>
<th>Schrödinger solution method #2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average time, s</td>
<td>%</td>
</tr>
<tr>
<td>heterostructure</td>
<td>0.068</td>
<td>0.01</td>
</tr>
<tr>
<td>condbandpotential</td>
<td>0.305</td>
<td>0.03</td>
</tr>
<tr>
<td>internal_field</td>
<td>0.049</td>
<td>0.00</td>
</tr>
<tr>
<td>find_field_effects...</td>
<td>1.049</td>
<td>0.09</td>
</tr>
<tr>
<td>solve_schroedinger</td>
<td>2.045</td>
<td>0.18</td>
</tr>
<tr>
<td>self_consistent...⋆</td>
<td>1130.813</td>
<td>99.67</td>
</tr>
<tr>
<td>wfplot</td>
<td>0.083</td>
<td>0.01</td>
</tr>
<tr>
<td>gain</td>
<td>0.142</td>
<td>0.01</td>
</tr>
<tr>
<td>current</td>
<td>0.003</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1134.557</strong></td>
<td></td>
</tr>
</tbody>
</table>

* Cumulative time over unlimited number of periods until the convergence of electron temperature $T_e$.

According to Fig. 4.1, the script called poisson_schroedinger should be analysed as well. The results for the time consumption of poisson_schroedinger are given in Table 4.2. According to this, the most inefficient module is solve_schroedinger in the case of Schrödinger solution method #2, and find_scattering_rates in the case of method #1. Opportunities for optimisation of other modules were not analysed.

4.2 Improvement of the Schrödinger solver

The most “time-wasting” part of the software is solving Schrödinger (script solve_schroedinger in Fig. 4.2) in the case of method #2. The method itself is realised in script fwwf (see Fig. 4.3). This is an eigenvalue problem (EVP) to find energy eigenvalues and corresponding wave functions of electrons in the system. Method #2 is one of the preferred methods QCLSIM uses for solving the Schrödinger equation, as it takes band nonparabolicity into account (see details in [31]). Although the method is very precise, it is also very time consuming. According to eq. (14) in [31], there is a matrix equation containing a large sparse matrix to be solved for EVP. All the calculations were originally performed using a dense matrix EVP solver from the LAPACK library. As Cooper has shown in [31], the linearised EVP
Table 4.2: Analysis of time consumption per sub-programs in `poisson_schroedinger`. The number of runs for `poisson_schroedinger` was 11 for method #1 and 17 for method #2. Sub-programs ran 8 times more cycles (see Fig. 4.2).

<table>
<thead>
<tr>
<th>Sub-program of poisson_schroedinger</th>
<th>Schrödinger solution method #1</th>
<th>Schrödinger solution method #2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average time, s</td>
<td>%</td>
</tr>
<tr>
<td>densityinput</td>
<td>0.026</td>
<td>0.22</td>
</tr>
<tr>
<td>find_fermi</td>
<td>0.062</td>
<td>0.52</td>
</tr>
<tr>
<td>find_scattering_rates</td>
<td>8.701</td>
<td>72.57</td>
</tr>
<tr>
<td>pop</td>
<td>0.044</td>
<td>0.37</td>
</tr>
<tr>
<td>find_fermi</td>
<td>0.046</td>
<td>0.38</td>
</tr>
<tr>
<td>find_field_effects...</td>
<td>1.007</td>
<td>8.40</td>
</tr>
<tr>
<td>chargedensity</td>
<td>0.073</td>
<td>0.61</td>
</tr>
<tr>
<td>gain</td>
<td>0.019</td>
<td>0.16</td>
</tr>
<tr>
<td>excrates</td>
<td>0.019</td>
<td>0.16</td>
</tr>
<tr>
<td>solve_schroedinger</td>
<td>1.985</td>
<td>16.56</td>
</tr>
<tr>
<td>delta</td>
<td>0.007</td>
<td>0.06</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>11.989</strong></td>
<td></td>
</tr>
</tbody>
</table>
could in principle be solved faster when using some custom library rather than using LAPACK.

According to [31], the EVP was solved using this matrix equation:

$$
\begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-A_3^{-1}A_0 & -A_3^{-1}A_1 & -A_3^{-1}A_2
\end{bmatrix}
\begin{bmatrix}
\psi \\
E\psi \\
E^2\psi
\end{bmatrix} = E
\begin{bmatrix}
\psi \\
E\psi \\
E^2\psi
\end{bmatrix},
$$

(4.1)

where $0$ is the matrix containing zeros, $1$ is the unit matrix, $A_0$, $A_1$, $A_2$ and $A_3$ are matrices containing kinetic and potential energies of electrons and applied fields and are described in [31], $E$ is the vector for energy eigenvalues of electrons, and $\psi$ is the matrix of corresponding wave functions. A more general representation of this equation is:

$$A\psi = E\psi,$$

(4.2)

where $A$ is the left-most matrix in eq. (4.1). Technically, matrix $A$ is equal to:

$$A =
\begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0
\end{bmatrix},
$$

where $a_i$, $b_i$, $c_i$, $d_i$, $e_i$, $f_i$, and $g_i$ are nonzero real coefficients.

As matrix $A$ is a sparse matrix, experiments with the ARPACK library (that supports sparse matrices) showed that simple swapping from general matrix methods to sparse matrix methods brings up converging problems. According to [32], the shift and invert spectral transformation should be done to improve convergence. This means that the original eq. (4.2) is to be transformed to

$$A =
\begin{bmatrix}
b_1 & c_1 & \cdots & 0 \\
a_1 & b_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & b_N
\end{bmatrix}
\begin{bmatrix}
e_1 & f_1 & \cdots & 0 \\
d_1 & e_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & e_N
\end{bmatrix}
\begin{bmatrix}
g_1 & 0 & \cdots & 0 \\
0 & g_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & g_N
\end{bmatrix},
$$
Table 4.3: Comparison of time spent (in seconds) on solving Schrödinger using three methods. Taylor is method #1, LAPACK is method #2, and ARPACK is the improved version of method #2.

<table>
<thead>
<tr>
<th>#</th>
<th>Taylor</th>
<th>LAPACK</th>
<th>ARPACK</th>
<th>Taylor</th>
<th>LAPACK</th>
<th>ARPACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.003</td>
<td>1.795</td>
<td>0.079</td>
<td>1.174</td>
<td>408.989</td>
<td>21.477</td>
</tr>
<tr>
<td>2</td>
<td>0.003</td>
<td>1.410</td>
<td>0.082</td>
<td>1.007</td>
<td>399.101</td>
<td>20.403</td>
</tr>
<tr>
<td>3</td>
<td>0.003</td>
<td>1.843</td>
<td>0.090</td>
<td>1.023</td>
<td>403.421</td>
<td>20.563</td>
</tr>
<tr>
<td>4</td>
<td>0.003</td>
<td>1.582</td>
<td>0.080</td>
<td>1.097</td>
<td>404.587</td>
<td>20.496</td>
</tr>
<tr>
<td>5</td>
<td>0.004</td>
<td>1.916</td>
<td>0.085</td>
<td>0.957</td>
<td>379.871</td>
<td>20.430</td>
</tr>
<tr>
<td>6</td>
<td>0.004</td>
<td>1.422</td>
<td>0.077</td>
<td>0.984</td>
<td>375.256</td>
<td>20.656</td>
</tr>
<tr>
<td>7</td>
<td>0.008</td>
<td>1.773</td>
<td>0.078</td>
<td>1.216</td>
<td>384.427</td>
<td>21.125</td>
</tr>
<tr>
<td>8</td>
<td>0.004</td>
<td>1.923</td>
<td>0.077</td>
<td>1.167</td>
<td>396.451</td>
<td>21.509</td>
</tr>
<tr>
<td>9</td>
<td>0.004</td>
<td>1.800</td>
<td>0.082</td>
<td>0.974</td>
<td>401.195</td>
<td>22.575</td>
</tr>
<tr>
<td>10</td>
<td>0.016</td>
<td>2.119</td>
<td>0.075</td>
<td>1.429</td>
<td>370.507</td>
<td>23.671</td>
</tr>
<tr>
<td>AVG</td>
<td>0.005</td>
<td>1.758</td>
<td>0.081</td>
<td>1.103</td>
<td>392.381</td>
<td>21.291</td>
</tr>
</tbody>
</table>

\[(A - \sigma)^{-1} \psi = E_\nu \psi\]

where \(\sigma \neq E\) is the shift in absolute values for elements of \(A\), and \(E_\nu = \sigma + \frac{1}{E}\) is a vector containing shifted eigenvalues. As ARPACK supported this kind of transformation, further experiments with this library still showed some convergence problems. By analysing and changing the magnitude of matrix \(A\) elements, the converging was improved. The magnitude of matrix \(A\) elements is around \(10^{-58}\), which is relatively small. By normalising the matrix, setting \(1\) as the average value of matrix \(A\), the converging problems disappeared. It has been mentioned that ARPACK works better with eigenvalues of a large magnitude [32]. Experiments showed much smaller calculation times when using the ARPACK library in shifted and inverted spectral transformation mode.

Table 4.3 and the corresponding Fig. (4.7) show time spent on two calculation series of the Schrödinger equation with two structures. Time saved when using the sparse matrix solver rather than general matrix solver is around \(\frac{1.758s}{0.081s} = 21.7\), with structure #1 and \(\frac{392.381s}{21.291s} = 18.4\) times with structure #2, giving approx. 20 times the average success.

### 4.3 Improvements due to the reuse of data

During the simulation process, many parameters are calculated over and over, but could be reused instead. In Fig. 4.8 you can see an example of subband populations calculated for a QCL for different electric fields. To
get characteristic curves as in this figure, the whole process, as seen in Fig. 4.1, needs to be repeated dozens or hundreds of times for each electrical field and lattice temperature value. We cannot directly use any data from previous calculations, but we could use previous data as an initial guess. The process in Fig. 4.2 contains an iterative part, which means that during the calculation some values are converging from their initial states to the values we are seeking. With initial values closer to the expected results we could decrease the number of iterations required.

If we could already leave out a few steps in the iterative process, we could gain in calculation time. A calculation on structure #1 with a few more cycles was made (see Fig. 4.9, $E = 18.7 \text{kV/cm, } T = 77 \text{K}$). The first calculation made was simply a calculation of subband populations from their default initial values, which were all equal to each other (each subband was initially equipopulated). The second calculation was purely to reuse data of the subband population from a previously completed calculation, where the electric field was $18.6 \text{kV/cm}$ (The difference from the current calculation is $0.1 \text{kV/cm}$). According to this, the initial value of electron population is pretty close to the convergence value later. In spite of that, the value starts fluctuating straight from the first iterations and merges quickly with the values from the original calculation, giving no direct positive impact on the speed of convergence.

The next calculation with data reuse was to merge potential-related data
together with population data into the initial data. Due to the field effects, the potential of the QCL structure is slightly changed during each iteration (stage 4 in Fig. 4.2). Incorporating the potential data could improve the efficiency of data reuse. As can be seen from Fig. 4.9, the effect is small but notable. The solid line shows that the number of subband populations remains close to the converging value for quite a few steps from the beginning, then suddenly jumps remarkably to a higher value, coming then back again close to the converging value. This experiment showed better results, but it is still too unstable to allow for a decrease in the number of iterations to be performed per one calculation.

### 4.4 Integration improvements

When analysing scattering rate calculations in the script called `find_scattering_rates`, a numerical issue was found. Numerical calculations may become very difficult if there are complex integrations during the simulation process. During the calculation of acoustic phonon scattering rate (script `srac` in 4.5), the following integration needs to be calculated numerically (see eq. (3.7) in this work, as well as eq. (10.186) in [3]).

According to this equation, $\alpha_1$ and $\alpha_2$ can cause integrals to become infinite, if $\alpha_1 = \alpha_2$. This equation can be solved analytically without any
Figure 4.9: Using population data from other calculations performed under similar conditions as the initial values provides a notable effect in convergence of subband populations. In this calculation, the initial number of populations per each subband was $1.657467 \cdot 10^{14} \frac{1}{m^2}$. This value was intentionally left out of this figure to give a better overview of what happens after the conversion process has been started.

This work has revealed that leaving out only these integration steps where $\alpha_1 = \alpha_2$ still causes roughness in scattering rate spectra of acoustic phonons (see the dotted lines in Fig. 4.10). In analysing the scattering rate’s exact dependence on $\phi$, it was seen that $\phi$ values near the $\alpha_1 = \alpha_2$ condition cause $\frac{1}{\tau_i}$ to grow unexpectedly high. Mathematically, it is clear that if the denominator approaches zero, the result approaches infinity. The physical meaning represent some specific angles between the initial and final momentum states of the carrier that are causing unexpectedly high scattering rates and should thus be avoided. According to the numerical experiments, the ‘smoothness’ of the scattering rate was improved, leaving out not only these integration steps where $\alpha_1 = \alpha_2$, but also neighbouring steps (as seen in Fig. 4.11).

In eq. (3.7) it can be seen that the overall value of the integral does not change if we replace integral $\int_0^{2\pi} d\phi$ with integral $2\int_0^\pi d\phi$. Doing this will make the solution process of this integration approximately 2 times faster.

Modified acoustic phonon calculation showed that the roughness in scattering rate spectra of acoustic phonons decreased, making the code more precise (see the solid lines in Fig. 4.10).
Figure 4.10: Small roughnesses in the scattering rate spectra of acoustic phonons [30]

Figure 4.11: Integration can cause problems...
4.5 Conclusion

In this chapter, three main ideas regarding how to optimise the QWWAD software tool were presented. The improved Schrödinger solver, which handles nonparabolicity, is now 20 times faster than it was before. Finding new values for initial guesses of subband calculations showed a slight opportunity for improvement in terms of the number of iterative cycles, which could be reduced. This approach needs further research. The third idea was related to numerical integrations that could be performed more precisely and faster. An example calculation showed smoother results in acoustic phonon scattering rates.
Conclusion

The main objective formulated initially was to analyse, develop and optimise the models and algorithms used for quantum well heterostructures. As a result an alternative source and relevant algorithms were investigated to produce incoherent spontaneous terahertz emission using lateral electron transport. In conjunction with this many algorithms were improved (incl. Schrödinger equation solver) to optimise the overall simulation process. Five articles related to this thesis have been published. One manuscript is awaiting acceptance [30].

The main results of this thesis are listed in the following:

1. According to Ch. 3, laterally pumped GaAs/AlGaAs quantum wells can be considered potential coherent broadband sources of terahertz radiation. The in-plane field accelerates electrons, heating them and increasing the population inversion. In the range around 7 THz, the power of the radiation rose over the black body radiation even at room temperature 300 K. The dependence between the lateral electric field and emitted power is shown in [III]. It has been shown that the population inversion takes place even near a temperature of 400 K [29]. QCLs (the main coherent terahertz sources) need low lattice temperature (under 200 K @ 2.85 THz [2]); therefore, any success in developing devices that work at room temperature is very important.

2. Ch. 2, [I] and [29] present a new method implemented to generate the digitised quasi-parabolic quantum wells, seeking equal spacings between energy eigenvalues. For 7 THz emission frequency, a root-mean-square deviation of energy spacings below 4% was achieved, which is enough for the practical approximation of parabolic quantum well.

3. According to Ch. 2 and [II], the most reasonable method to solve the Schrödinger equation numerically is the standard linear algebra solver (for example, LAPACK and ARPACK). Solving difference equations numerically showed remarkably worse results.
4. According to Ch. 4 and [30], ARPACK shift-inverted sparse matrix methods gave approximately 20 times faster algorithms than LAPACK dense matrix methods in solving the Schrödinger equation, taking band nonparabolicity into account. The matrix equation composed for this solution contains mostly zeros, which makes it reasonable to use sparse matrix methods instead of methods for general dense matrices.

5. According to Ch. 4 and [30], the amount of time the QCL simulation algorithm runs depends on the initial guess values for the subband population. Experiments showed a slight opportunity for improvement in the number of iterative cycles that could be reduced. This approach needs further research.

6. Several calculation modules originally taken from [3] (QWWAD) were re-implemented, and a few new modules were built. The list of modified and newly implemented calculation modules is given in Table C1.

7. An overview of the advances in terahertz technology was written and published as [IV].

In the future, the algorithms to simulate semiconductor quantum well heterostructures should be optimised further to achieve a fast simulation platform without significant calculation times. Although the increase in computing power will continue for a long time, the methods of calculation still need to be improved too. New models should also be built to achieve better compliance with the real physical environment.

A user interface that is easy to use should be built to simplify the learning process for using these simulation tools. It would be ideal for the interface to be accessible via web page (i.e. as an applet), providing very broad access to the simulation methods and broadening the scientific opportunities in this area.
Table C1: The list of QWWAD calculation modules modified and added.

<table>
<thead>
<tr>
<th>Code</th>
<th>New or modified</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>fwf</td>
<td>modified</td>
<td>Calculation of Schrödinger equation using nonparabolic approach. The algorithm was improved, which is now 20 times faster.</td>
</tr>
<tr>
<td>sbp</td>
<td>modified</td>
<td>Calculation of quasi-Fermi energies for each subband. The algorithm was improved to allow nonzero electric fields.</td>
</tr>
<tr>
<td>srac</td>
<td>modified</td>
<td>Calculation of acoustic scattering rates. The precision of the algorithm was improved.</td>
</tr>
<tr>
<td>srmpr</td>
<td>modified</td>
<td>Calculation of mean scattering rates. The Fermi-Dirac distribution function was replaced by the shifted FD function. Nonzero electric field support was also added.</td>
</tr>
<tr>
<td>srrad</td>
<td>modified</td>
<td>Calculation of spontaneous radiative lifetimes. The method in the original version was improved to take laterally pumped electron transport into account.</td>
</tr>
<tr>
<td>ND2ni</td>
<td>new</td>
<td>Calculation of subband populations using rate equations.</td>
</tr>
<tr>
<td>radpow</td>
<td>new</td>
<td>Calculation of generated radiation power and spectrum.</td>
</tr>
<tr>
<td>finite_difference_method</td>
<td>new</td>
<td>Calculation of Schrödinger equation using standard matrix eigenvalue solver LAPACK.</td>
</tr>
</tbody>
</table>
References


The author would like to thank supervisors Dr. Andres Udal and Prof. Emer. Enn Velmre for the 10-year-long fruitful cooperation in this field. The author thank also Prof. Toomas Rang and other researchers from Thomas Johann Seebeck Department of Electronics in Tallinn University of Technology for the support during my studies.

Special thanks to Prof. Paul Harrison from Leeds University, who gave the author several opportunities for cooperation with his research group, including a four month period in Leeds University, and Dr. Alexander Valavanis for an exclusive access to the live repository of the source code of QWWAD software tool. The author would also like to thank the researchers from the School of Electronic and Electrical Engineering, University of Leeds (Dr. Zoran Ikonic, Dr. Craig Evans, Dr. Leon Lever etc) for all kind of support during the cooperation.
Abstract

Improvement and optimisation of simulation software for terahertz range radiation sources based on quantum well heterostructures

This thesis addresses the field of semiconductor quantum electronics and is based mainly on tasks to solve the Schrödinger equation and model electron transport in GaAs/AlGaAs multibarrier heterostructures. An overview of the latest achievements in the terahertz region of the electromagnetic spectrum along with basics of the main principles for semiconductor emitters is given in the first chapter. Numerical solution methods to solve a 1D time-independent Schrödinger equation are compared and tested in Ch. 2. It was found that the most optimal way to solve the Schrödinger equation is to use standard linear algebra solvers for eigenvalue problems. Calculations are performed for digitally graded parabolic quantum wells and results are presented in 3 separate publications. The methodology for calculating the emission characteristics for laterally pumped quantum well heterostructure-based emitters is given in the third chapter. The effect of lateral electric field is explored and relevant methods presented. The results show that a real emitter can be built using only one layer of GaAs bordered by Al\textsubscript{0.42}Ga\textsubscript{0.58}As layers. For the 7 THz emitter the overall emitted power rises higher than black body radiation. There is one publication on this topic. Optimisation strategies to improve the speed and accuracy of methods in QWWAD QCL simulation tools are handled in chapter 4. The results show an algorithm which is approximately 20 times faster in Schrödinger equation when using sparse matrix methods in the ARPACK library. There is one paper published and one paper waiting to be published on this topic.

Key words: Schrödinger equation, shifted Fermi-Dirac distribution function, digitised parabolic quantum well, lateral transport, terahertz emitter, quantum cascade laser
Kokkuvõte

Kvantaukudega heterostruktuuridel põhinevate terahertskiirguri-te simulatsioonitarkvara täiendamine ja optimeerimine

Töö käsitleb pooljuht-kvantelektroonika valdkonda ja baseerub peamiselt Schrödingeri võrrandi lahendamise ülesannetele ja elektroni transpordi modelleerimisele GaAs/AlGaAs mitmikbarjääridega heterostruktuurides. Töö esimeses osas antakse ülevaade elektromagnetkiirguse spektri terahertzpiirkonna viimastest saavutustest koos pooljuhtkiirgurite aluste ja tööpõhimõtetega. Töö teises osas on võrreltud ja proovitud ühemõõtmelise ajast sõltumatu Schrödingeri võrrandi numbrilisi lahendusmeetodeid. Optimaalseimaks viisiks Schrödingeri võrrandi lahendamiseks osutus standardne lineaaralgebra lahendusmeetod omaväärtusprobleemi jaoks. On tehtud arvutusi digitaalselt tasandatud paraboolsete kvantaukude jaoks ja tulemusted on avaldatud 3 eraldi artiklis. Töö kolmandas osas esitatakse lateraalselt ergastatud kvantauk-heterostruktuuridel põhinevate kiirgurite emissioonikarakteristikute arvutamistemoodika. Uuritakse lateraalse elektrivilja mõju ja tutvustatakse seonduvaid meetodeid. Tulemusted näitavad, et ühel GaAs kihil, mis on ümbrisestud Al_{0.42}Ga_{0.58}As kihtidega, põhinevat kiirgurit on reaalselt võimalik ehitada. 7 THz kiirguri jaoks ületab kiirata tõenisust. Sellal teemal on avaldatud üks artikkel. Töö neljas osa käsitleb QWWAD QCL simulatsioonihendite meetodite optimeerimist kiiruse ja täpsuse osas. Tulemusted näitavad umbes 20 korda kiiremat Schrödingeri lahendamise algoritmi, kui kasutada ARPACK teegi hõredate maatriksite meetodeid. Sellal teemal on avaldatud üks publikatsioon, veel üks artikkel on avaldamise ootel.

Võtmesõnad: Schrödingeri võrrand, nihutatud Fermi-Diraci jaotusfunktioon, digitaliseeritud paraboolne kvantauk, lateraalne transport, terahertskiirgur, kvantkaskaadlaser
Appendices
Numerical investigation of digitised parabolic quantum wells for terahertz AlGaAs/GaAs structures

R. Reeder, A. Udal, E. Velmre, P. Harrison

1Department of Electronics, TTU, Ehitajate tee 5, 19086 Tallinn, Estonia, e-mails: reeno@cyber.ee, audal@va.ttu.ee, evelmre@ttu.ee
2School of Electronic & Electrical Engineering, University of Leeds, LS2 9JT, U.K., e-mail: p.harrison@leeds.ac.uk

ABSTRACT: The calculations of quasi-parabolic quantum wells (QW) with smooth and digitised potential distributions on GaAs/Al0.42Ga0.58As system for 1–7 THz frequencies are accomplished by numerical one-dimensional Schrödinger equation solvers. It is shown that for energy eigenvalues equal spacing in smooth QWs the 18% quartic addition to parabolic potential is needed to compensate the electron effective mass increase in AlxGa1-xAs as x → 0.42. The optimisation of digitised multibarrier distributions revealed existence of two types of optimal solutions. At that the second type with periodically filled centre yielded low rms-deviations below 5% for energy spacings in frequency range 4÷7 THz. For frequencies below 2 THz the accuracy of digitised approach became insufficient to guarantee reasonably equally spaced energy levels.

1. Introduction

The quantum well (QW) with parabolic potential distribution is a simple classical quantum physics task but energy levels equal spacing in it opens very attractive possibilities for optoelectronic applications which are realised on semiconductor heterostructures. The most investigated and used GaAs/AlxGa1-xAs system gives possibilities to create QWs with deepness of 0.3÷0.4 eV range for electrons. This suites well for terahertz region 1÷10 THz applications (∆E = 4.135 ÷ 41.35 meV, λ = 30 ÷ 300 µm, QW sizes ca 50 ÷ 500 nm).

Generally rectangular QWs with abrupt GaAs/AlxGa1-xAs interfaces are the easiest to fabricate but with molecular-beam epitaxy (MBE) growth method the engineering of various potential profiles became feasible since 1980-ies [1]. To realise the parabolic quantum wells (PQW), the gradual changing of the material composition x [2] or use of the digital grading (DG) [1,3,4] are possible. The DG method bases on alternating deposition of thin layers of two fixed compositions GaAs and AlxGa1-xAs so that in average the influence close to smooth potential is obtained. At that the digitised barrier/well layer widths may be reduced to one crystalline monolayer (1 m.l. ≈ 0.283 nm), see e.g. [5,6].

Below we report the one-dimensional Schrödinger equation numerical solution results in order to evaluate feasibility of energy eigenvalues equal spacing with accuracy ± 5% on GaAs/Al0.42Ga0.58As system for frequencies 1÷7 THz. The smooth and digitised potential (layer thickness 1 m.l. and 2 m.l.) cases are considered. Figure 1 presents one possible application – a broadband terahertz radiation source.

2. Models and numerical algorithms

The basis for calculations is the time-independent 1D Schrödinger equation:

\[ -\hbar^2 \frac{\partial^2}{\partial z^2} \left( \frac{1}{m^*(z)} \right) \psi(z) + V(z) \psi(z) = E \psi(z) \]

where a modified kinetic energy term is used as commonly accepted in variable effective mass semiconductor QW calculations [7, p.44]. The condition for defining of energy eigenvalues is the vanishing wave functions \( \psi \to 0 \) at active area boundaries.

The two numerical algorithms were used independently and/or in a combined way: a) the energy-trying shooting method with open boundaries [7] and b) the Newton method solver realising the simultaneous calculation of wave functions and energy values with zero boundary conditions. For both methods the usual three-point discretisation in space was applied for kinetic energy term [7]. The shooting method was found reliable to scan over the energies interval to obtain approximate solutions. Using those precalculated solutions, the Newton method exhibited superior convergence speed (usually 3-4 iterations with efficient 3-diagonal bordered
matrix internal solver). The two necessary starting \( \psi \) values in QW centre for shooting method were specified from symmetric and antisymmetric wave function conditions \([7]\). The shooting method’s idea of stepping over space and energy is explained by Figure 2.

Figure 2 Algorithm of shooting method

The band gap and electron effective mass linear dependences on Al percentage \( x \) were adopted from \([7, p.431]\) what gave QW deepness \( V_{\text{max}} = 0.3509 \text{eV} \) for \( x_{\text{max}} = 0.42 \) and increase of \( m^* \) from 0.067 to 0.102 if \( x \) increases from 0 to 0.42. For digitised potential barriers and wells the basic step 1 monolayer (0.2825 nm) was taken and rectangular potential shape was assumed. The typical spatial discretisation step was 0.01 nm.

3. Results for smooth diffused QWs

Figure 3 shows the investigated smooth potential profiles and variation of effective mass in the case of 7 THz PQW. Figure 4 gives the calculated energy levels and level spacings for highest and lowest analysed frequency. Figure 5 shows the form of calculated 12 wave functions in 7 THz PQW. To specify the potential distributions, the wanted energy spacings \( \Delta E = (28.95 \text{meV}) (f/7 \text{THz}) \) and relevant ideal PQW widths \( l_{QW} = (61.72 \text{nm})(7 \text{THz}/f) \) (at \( E = V_{\text{max}} \) level for unlimited parabola; \( m^* = 0.067 \)) were used as the backing values. Figure 4 shows that the limited height of QW excludes only one level out of 5% criterion. In contrast to that the variable effective mass causes significant decrease of higher energy spacings. However, introduction of 18% quartic addition to parabolic potential

\[
V(z) = V_{\text{max}} (z / 0.5 l_{QW})^2 + 0.18 V_{\text{max}} (z / 0.5 l_{QW})^4
\]

nearly entirely compensates this effect.

Figure 3 Specified potential distributions and electron effective mass distribution in parabolic GaAs/Al_{0.42}Ga_{0.58}As quantum well, 7 THz case shown

Figure 4 Calculated energy spacing versus energy in parabolic quantum wells: 7 THz above and 1 THz below

Figure 5 Calculated wave functions (in arbitrary units, raised by energy eigenvalues) and parabolic potential (with 18% \( z^4 \)-term) in 7 THz parabolic quantum well
4. Results for digitised quantum wells

The digitised QW is actually a complex multibarrier/multiwell system what is rather demanding task for calculation and especially for optimisation to obtain equal energy eigenvalues spacing. Since every pillar-like potential barrier causes abrupt inversion of wave function 2nd derivative (curvature) it is expected that to maintain wave function picture similar to Fig. 5, the barrier width must be kept minimal, i.e. 1 monolayer. This is confirmed by Table 1 where optimisation results for 2 m.l. barrier/well step are clearly worse than for 1 m.l. step. All other results below correspond to step 1 m.l.

The search of optimal digitised potential distributions was started from 7 THz, where numbers of energy levels (11÷12) and potential pillars (ca 30, max ≈200) are the lowest. Initial distribution was specified from condition that averaged digital potential corresponds to the smooth parabola. In optimisation the barrier/well inversions and barrier shifts (single and group) were tried to minimise the root mean square (RMS) deviation of energy spacings ( in respect of wanted exact $\Delta E$ ). The optimisation procedure revealed the existence of two types of optimal distributions:

1) Type 1 “classical parabolic QW like” with empty QW centre (see below Figs. 6 and 9) and low first energy level $E_1 = (0.15 \div 0.35)\Delta E$;
2) Type 2 “superlattice centre type” with nearly periodically (period 12÷13 monolayers) filled by single pillars centre (see below Figs. 7 and 10) and with high frequency-independent first energy level $E_1 = (28 \div 42)$ meV.

Table 1 Comparison of minimised RMS-deviations of energy level spacing for the barrier/well width technological steps of 1 and 2 monolayers (for 7 THz)

<table>
<thead>
<tr>
<th>Optimised digital distribution type</th>
<th>RMS-deviation of $\Delta E$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 monolayer</td>
</tr>
<tr>
<td>Type 1 “Empty centre”</td>
<td>8.0%</td>
</tr>
<tr>
<td>Type 2 “Periodically filled centre”</td>
<td>3.3%</td>
</tr>
</tbody>
</table>

Figure 6 Optimised digital potential distribution of type 1 and respective wave functions (in arb. units), rised by energy eigenvalues in 7 THz quantum well

The optimised type 2 distributions gave approximately twice lower $\Delta E$ RMS-deviations than type 1 (see Figs. 8 and 11, also Table 1). The wave functions for type 2 potential, especially the lower ones, obtained specific curved fragments with potential pillars period (Fig. 7). Note that the $\Delta E$ maximum deviations were typically 2÷2.5 times higher than the RMS-deviations from Fig.8.

Figure 8 Spacing of energy eigenvalues for 7 THz (a), 5 THz (b), 3 THz (c), 2 THz (d) digitised quantum wells

Figure 7 Optimised digital potential distribution of type 2 and respective wave functions (in arbitrary units), rised by energy eigenvalues in 7 THz quantum well
Figures 9 and 10 compare at four frequencies type 1 and type 2 optimised digitised potential distributions. Note that at all frequencies the period of central potential pillars in type 2 is 12÷13 monolayers. Figure 11 summarises the achieved $\Delta E$ RMS-deviations from performed optimisation calculations. 

![Figure 9](image1.png)

**Figure 9** Comparison of optimised potential distributions type 1 for frequencies 7 THz (a), 5 THz (b), 3 THz (c) and 2 THz (d)

![Figure 10](image2.png)

**Figure 10** Comparison of optimised potential distributions type 2 for frequencies 7 THz (a), 5 THz (b), 3 THz (c) and 2 THz (d)

![Figure 11](image3.png)

**Figure 11** Achieved minimised RMS-deviations of energy level spacing $\Delta E$ at different operation frequencies for digital potential distributions of type 1 and 2

### 5. Conclusions

General results from Fig.11 show that the goal 5% is feasible for frequencies 4÷7 THz if we consider RMS-deviations. However, the maximum $\Delta E$ deviations which were ca 2÷2.5 times higher, would yield only 8÷10% for same frequencies. The main reason, why at low frequencies the achieved accuracy is lower, is the increasing number of energy levels for what all is difficult to find satisfying locations of potential pillars. It should be mentioned that in present preliminary study we could not get reliable digitised potential distribution for the lowest frequency 1 THz. This is because of computer time growth as $-1/f^3$ per one optimisation sweep (number of energy levels, potential pillars and spatial steps all grow as $-1/f$) and because of breaking of normal optimisation process due to inevitable appearance of some very closely located energy levels in the case of greater number of levels and potential pillars.

If only part of energy range could be selected out by importance, the general results might be essentially better. In practical device design could be considered use of lower Al percentage (lower QW depth and smaller number of energy levels) at lower frequencies.

To estimate practical feasibility of current results, in the following the sensitivity of $\Delta E$ deviations to digitised potential small changes must be studied. Also the barrier periodity results observed here need deeper theoretical generalisation.

### Acknowledgement

This work has been supported by Estonian Science Foundation grant no. 5911.

### References


Comparison of methods for solving
the Schrödinger equation for multiquantum
well heterostructure applications

Andres Udal\textsuperscript{a}, Reeno Reeder\textsuperscript{a}, Enn Velmre\textsuperscript{a} and Paul Harrison\textsuperscript{b}

\textsuperscript{a} Department of Electronics, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia; audal@va.ttu.ee, reeno@cyber.ee, evelmre@ttu.ee

\textsuperscript{b} School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT, United Kingdom, p.harrison@leeds.ac.uk

Received 23 May 2006, in revised form 7 July 2006

Abstract. Direct numerical approaches for the solution of the time-independent one-dimensional Schrödinger equation are discussed. Applications to multiquantum well (MQW) semiconductor heterostructure potentials need linear dependence of the computer time $t_{comp}$ on the number of spatial grid points $N$. Although acknowledged as a very effective Fourier grid Hamiltonian (FGH) method, it has cubic dependence on the number of spatial grid points, i.e., $t_{comp} \sim N^3$, which limits its use to problems with a complexity of $N \leq 1000$. A simple straightforward shooting method (ShM), which is based on trial stepping over the coordinate and energy, has the necessary $t_{comp} \sim N$ dependence with moderate energy convergence efficiency but the recommended symmetry preconditions and the not very clearly defined external boundaries make its application inconvenient. This paper offers a new reliable and effective energy and wave function coupled solution (EWC) method with a Newton iteration scheme and an internal bordered tridiagonal matrix solver. The method has a linear $t_{comp} \sim N$ dependence and may be applied to arbitrary potential energy distribution tasks with complexity up to $N = 10^5$ and beyond. Zero or cyclic boundary conditions may be specified for the wave function. For versatile MQW tasks the combined use of ShM and EWC is illustrated. Detailed accuracy and computer time comparisons show that the combined ShM+EWC method is three orders of magnitude more effective than the FGH method.

Key words: multiquantum well structures, Schrödinger equation, bound states, numerical methods, energy, wave function.

1. INTRODUCTION

In spite of the fact that Erwin Schrödinger formulated his famous equation 80 years ago in 1926 and that for over 70 years scientists have proposed various analytical and numerical methods for the solution of this central quantum
mechanics equation, approaches even to one-dimensional solutions are still a subject of debate. This is confirmed by the continuing appearance of new publications in this field \([1-6]\). Since 1990s, one of the driving forces in this field has been physical chemistry and its applications. The second spur of motivation has come from the extremely wide application area in semiconductor heterostructures with quantum wells, wires and dots \([7]\). Although nowadays the available sophisticated \textit{ab-initio} software tools already make quite realistic three-dimensional calculations possible, almost every research task needs an estimation of static bound states in a one-dimensional (1D) approximation either in spherical or rectangular coordinates. The last case is more typical for semiconductor heterostructures where the calculation of 1D bound states for complex multi-barrier quantum well systems like, e.g. in quantum cascade lasers \([8]\) or digitized quasi-parabolic quantum wells \([9]\), may be a rather time-consuming subtask.

In applications in physical chemistry, the potentials in the atomic subnanometer scale are rather smooth and relatively small number \((N \leq 100)\) of spatial grid points may be sufficient to obtain accurate results. In contrast to that, MQW structures with great numbers of relatively abrupt potential steps over the 10–1000 nm spatial scale may need spatial grid sizes over \(N > 10^4\) to achieve acceptable results \([7,9]\). This means that for the analysis and even more for the optimization of MQW structures the methods of solving the Schrödinger equation must consume computer time no more than proportionally to \(N\).

The FGH method \([1-3]\) has been declared to be very effective and the simplest method for the calculation of bound states from the time-independent 1D Schrödinger equation. It can take as input an arbitrary potential distribution, but its computer execution time \(t_{\text{comp}}\) scales as \(N^3\). The use of the Fast Fourier Transform in the construction of the Hamiltonian matrix may reduce the last number to \(N^2 \times \ln N\) but not more \([2]\). The internal algebraic solver of the FGH method uses standard procedures from the EISPACK computer package \([10]\).

Actually, the most elementary solution method with the required linear dependence of the computer time on the grid size \((t_{\text{comp}} \sim N)\) is the shooting method \([7]\), which finds eigenvalues of the bound state energy by using a trial procedure, based on the condition that an iteration over the spatial coordinate from the centre of the active area into the surrounding barriers must yield a vanishing wave function \(\Psi \to 0\). The ShM needs two initial values of the wave function to start the iteration over the spatial coordinate. Those two values may be correctly defined for symmetric potentials but in the general case it needs an approximate auxiliary algorithm \([7]\). Another problem, troubling the application of the ShM, is associated with the exponentially growing components of the wave function in the outer barrier regions, which can impede the detection of the theoretical boundary condition \(\Psi \to 0\).

In the present work we propose an effective iterative coupled energy and wave function method for solving the time-independent one-dimensional Schrödinger equation, which possesses linear \(t_{\text{comp}} \sim N\) dependence and may be applied to arbitrary potential energy distributions without any symmetry restrictions. The
method uses an iterative approach in the Newton method and its internal algebraic task involves the solution of a linear system, represented by a bordered tri-diagonal matrix; zero or cyclic boundary conditions may be specified. In the case of arbitrarily complex MQW problems the combined ShM+EWC approach is necessary when approximate wave functions and estimations for the energy eigenvalues for the EWC method are supplied by the ShM. Considering formation of the Jacobi matrix in the Newton method by zero boundary conditions, the EWC method is similar to the known relaxational approach \[^4\]. As one can see below, direct comparison of FGH and ShM+EWC methods reveals that the latter may be over three orders of magnitudes more effective than the FGH method for tasks with \(N > 1000\). Some results of application of the present work are published in \[^9\].

In the present study we discuss neither semi-analytical approaches with specific application areas like \[^6\] nor the new original “random trial” approaches \[^7\].

2. METHODS OF THE NUMERICAL SOLUTION

2.1. The Fourier grid Hamiltonian method

The Fourier grid Hamiltonian method for numerical solution of the time-independent Schrödinger equation was introduced in 1989 by Marston and Balint-Kurti \[^1\]. Initially it was formulated for an odd number of coordinate grid points \(N\) in rectangular coordinates \[^1\]. Later it was modified for an even number of grid points and for spherical coordinates with the possibility of applying the Fast Fourier Transform to accelerate the formation of the Hamiltonian matrix \[^2,3\]. The theory behind the method is based on relating the potential energy at the \(N\) grid points with the kinetic energy in the momentum space via forward and reverse Fourier transforms between the coordinate and the momentum space. The \(N\times N\) symmetric matrix \(H\), obtained by discretization, has elements in the form of cosine sums. The task of calculating the bound state eigenenergies and eigenfunctions is thereby transformed to the task of finding eigenvalues and eigenvectors of the matrix \(H\). As suggested by the authors of the FGH method, this may be accomplished by standard subroutines such as the EISPACK package \[^10\]. The source code of the computer implementation of the method, FGHEVEN \[^2\], for an even number of nodes, is freely distributed via the internet \[^11\]. Subsequently, the method was extended to the three-dimensional case \[^12\]. In the present work we have realized the odd number FGH method in rectangular coordinates and in the SI unit system for testing, following reference \[^1\].

Assuming that the length of the calculation area \(L\) is divided into \(N\) steps, the discrete grid is defined as

\[
x_i = (i-1/2)\Delta x, \quad \Delta x = L/N, \quad i = 1, 2, ..., N.
\]

Equation (1), which defines nodes in the centre of every interval \(\Delta x\), differs somewhat from the original one \(x_i = i\Delta x\) \[^1\] but is more correct in the case of symmetric QW tasks.
Following [1], the spatial discretization and transformations performed yield a symmetric matrix $H$ of size $N \times N$

$$H_{ij} = \frac{2}{N} \sum_{i=1}^{n} \cos\left(\frac{2\pi(i - j)}{N}\right)T_j + V(x_j)\delta_{ij}, \quad n = (N-1)/2,$$

where $h = h/2\pi$ is the reduced Planck constant, $m$ is the electron rest mass and the Kronecker symbol $\delta_{ij}$ ensures that potential energy values at grid nodes $V(x_j)$ are only added to the main diagonal.

The EISPACK subroutines return $N$ eigenvalues and $N$ eigenfunctions with values at all $N$ grid nodes. The eigenvalues, which lie below the potential energy values at the solution area borders $V(0)$ and $V(L)$, may be interpreted as the bound state energies of the system [1]. The other eigenvalues may be interpreted as extraneous solutions, which unfortunately consume computer time for their calculation. On the other hand, the maximum number of eigenvalues is limited by the grid size $N$. As experience shows, in practical quantum well calculations sufficiently thick outer barrier layers with a sufficiently high potential $V$ must be included to ensure sufficient decay of wave functions and hence remove any uncertainty in the boundary conditions in the FGH method. It should be added that by selecting different subroutines from the EISPACK package, it is possible to calculate the energies without wave functions [11], which reduces the computation time approximately 1.5 times.

### 2.2. The shooting method

The shooting method is one of the simplest numerical algorithms and its key idea is to replace a boundary condition problem with multiple trial runs of a remarkably simple initial condition task. This is also the most straightforward method for solution of the time-independent Schrödinger equation if the varied trial parameter is the energy $E$ ([1], Chapter 3). The trial energy equals the bound state energy if the wave function vanishes $(\Psi \to 0)$ in the surrounding barriers when moving away from the QW area. In spite of the extreme simplicity, ShM has a linear dependence of the computer time on the grid node number $t_{\text{comp}} \sim N$, which makes it efficient for MQW tasks.

The time-independent 1D Schrödinger equation in its classical form, where kinetic energy is defined by the second spatial derivative, reads

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi}{dx^2} + V\Psi = E\Psi. \quad (3)$$

In the present study, which is focused on the effectiveness of numerical methods, we ignore the fact that in semiconductor heterostructures, a more sophisticated kinetic energy term [7]
\[
\frac{-\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x) \Psi = E \Psi,
\]

is recommended for tasks with a variable effective mass \( m(x) \).

To realize discrete stepping over the spatial coordinate, the second derivative in the Schrödinger equation (3) is replaced by a three-point discretization scheme

\[
\frac{\hbar^2}{2m} \left( \frac{\Psi_{i+1} - 2\Psi_i + \Psi_{i-1}}{\Delta x^2} \right) + V_i \Psi_i = E \Psi_i,
\]

where \( \Psi_i \) denotes \( \Psi(x_i) \) and \( V_i \equiv V(x_i) \).

To initialize the iteration in Eq. (5), two starting values of the wave function are needed. That is facilitated by the obvious fact that the Schrödinger equation (3) and the energy eigenvalues defined by it are insensitive to constant multipliers of \( \Psi \). That means that prior to the wave function normalization with the condition

\[
\int_{-\infty}^{\infty} \Psi^2(x) \, dx = 1,
\]

the scale of the wave function has no importance.

Thus for symmetrical QWs, the initial conditions in the QW centre \( x_c \) may be specified in the case of symmetrical wave functions by [7]

\[
\Psi(x_c) = 1, \quad \Psi(x_c + \Delta x) = 1 + \left( \frac{m \Delta x^2}{\hbar^2} \right) (V(x_c) - E),
\]

and for antisymmetrical wave functions by [7]

\[
\Psi(x_c) = 0, \quad \Psi(x_c + \Delta x) = 1.
\]

In the more general non-symmetrical case, the stepping must start from two adjacent grid points within one barrier, e.g. in the left barrier from coordinates \( x_{\text{left}} \) and \( x_{\text{left}} + \Delta x \):

\[
\Psi(x_{\text{left}}) = \epsilon, \quad \Psi(x_{\text{left}} + \Delta x) = \epsilon \exp(\kappa \Delta x),
\]

where \( \epsilon \) is a small (but finite) number and \( \kappa \) is the theoretical wave function exponential growth/decay constant, which follows directly from Eq. (3) for the constant barrier height \( V > E \) and energy eigenvalue guess \( E \) as [7]

\[
\kappa = \sqrt{2m(V - E) / \hbar^2}.
\]

The condition for energy eigenvalue detection is the vanishing wave function in the opposite external barrier: \( \Psi \to 0 \) as \( x \to -\infty \) ([7], p 75). Beside numerical problems, associated with the great range of numerical values of \( \Psi \), the advantage of considering only half of the structure in the symmetrical case is lost in the general non-symmetrical QW case.

One principal inconvenience with the numerical realization of the shooting method is associated with detecting the condition \( \Psi \to 0 \). The reason for this is
the fact that the Schrödinger equation (3) as a second order differential equation has solutions which contain both growing and decaying exponentials in a uniform potential barrier:

$$\Psi(x) = A \exp(-\kappa \Delta x) + B \exp(+\kappa \Delta x), \quad (9)$$

where the constant $\kappa$ is defined by Eq. (8). This theoretical form of the wave function in the outer barriers, surrounding the QW, means that every small change in the trial energy from the exact eigenvalues, as well discrepancies in arithmetic operations (rounding errors) are amplified as the stepping iteration proceeds into the barrier. Consequently, in a numerical realization of the ShM algorithm, instead of the theoretical condition $\Psi \rightarrow 0$ the energy value must be sought that changes the sign of the “tail” of the wave function ([7], p 77).

The idea of the shooting method of stepping over space and energy is explained in Fig. 1.

![Fig. 1. Algorithm of the shooting method. The internal cycle performs the stepping over space and two external cycles find the eigenvalues of the energy.](image-url)
2.3. The method of coupling the energy and wave functions

As seen above, realization of the shooting method demands a rather small amount of computations but the definition of the boundary conditions and the process of finding the exact energy eigenvalues need special care. Serious numerical problems may arise in the case of relatively thick barriers $(\kappa \gg 1)$, where exponential growth of the wave function follows from an inaccurately defined energy eigenvalue or discrepancies due to the limited number of significant figures available for the arithmetic operations. To avoid these problems, the EWC method was developed, which solves system of equations with clearly fixed boundary conditions simultaneously for the energy eigenvalue and wave function values across the spatial grid nodes. The 3-point scheme of spatial discretization used corresponds exactly to that of the shooting method in Eq. (5). The exact zero or cyclic boundary conditions for $\Psi$ for the calculations are fixed. The zero boundary conditions actually correspond to the assumption about infinite potential barriers on the external borders of the calculation area. To find the energy eigenvalue together with the values of the wave function on the grid nodes, an additional equation is necessary, besides the discretized Schrödinger equation, which is the normalization condition of the wave function (6). The latter states that the probability of finding an electron over the entire space of the calculation equals unity. With this additional condition the calculated wave functions from the EWC method are automatically normalized.

The unknown vector $Y$ of the EWC method contains $N$ components: the energy eigenvalue $E$ and wave function values $\Psi_i = \Psi(x_i)$ in nodes $i = 2, 3, \ldots, N$ of the grid:

$$x_i = (i - 1) \Delta x, \quad \Delta x = L/(N - 1), \quad i = 1, 2, \ldots, N,$$

$$Y = (E, \Psi_2, \Psi_3, \ldots, \Psi_N)^T,$$

where the superscript $T$ denotes transposition.

The value of the wave function in the 1st spatial node is not included in $Y$ since

$$\Psi_1 = \Psi_N$$

is assumed for both boundary condition types.

The zero boundary conditions are then specified simply as

$$\Psi_N = 0.$$ 

The cyclic boundary condition may be specified by the discrete Schrödinger equation (5) for the boundary node $N$. Taking into account the translational
symmetry, the node \( N \) is equivalent to node 1, node \( N+1 \) is equivalent to node 2 etc. Thus the necessary three “neighbour” wave function values in the boundary node \( N \) are \( \Psi_{N-1}, \Psi_{N}, \Psi_{2} \).

Thus the first non-linear equation of the EWC method system is the rewritten normalization condition of the wave function (6) in discrete form

\[
F(N, \Psi_{1}, ..., \Psi_{N}) \equiv \sum_{i=2}^{N} \Psi_{i}^{2} \Delta x - 1 = 0.
\]

The next necessary \( N-2 \) equations are the discrete Schrödinger equations (5) for the internal grid nodes \( i = 2, 3, ..., N-1 \):

\[
F(E, \Psi_{i-1}, \Psi_{i}, \Psi_{i+1}) \equiv \frac{\hbar^{2}}{2m} \left( \frac{\Psi_{i+1} - \Psi_{i}}{\Delta x} - \frac{\Psi_{i} - \Psi_{i-1}}{\Delta x} \right) \frac{1}{\Delta x} + (E - V_{i}) \Psi_{i} = 0, \quad (15)
\]

where for \( i = 2 \) according to Eq. (12) holds \( \Psi_{i-1} = \Psi_{N} \).

In the case of zero boundary conditions, the last equation of the system is

\[
F(N, \Psi_{N}) \equiv \Psi_{N} = 0. \quad (16)
\]

In the case of the cyclic boundary condition, the last equation of the system is similar to Eq. (15) with the replacement \( \Psi_{N+1} = \Psi_{2} \):

\[
F(N, \Psi_{N-1}, \Psi_{N}, \Psi_{2}) \equiv \frac{\hbar^{2}}{2m} \left( \frac{\Psi_{2} - \Psi_{N}}{\Delta x} - \frac{\Psi_{N} - \Psi_{N-1}}{\Delta x} \right) \frac{1}{\Delta x} + (E - V_{N}) \Psi_{N} = 0. \quad (17)
\]

The system (14)–(17) is non-linear as Eq. (14) contains squares of the values of the wave function and Eqs. (15) and (17) contain products of energy and wave functions. This system may be linearized and solved iteratively using the Newton method. For every iteration, the unknown vector \( Y \) may be written as

\[
Y = \tilde{Y} + \delta Y, \quad (18)
\]

\[
[\partial F / \partial Y] \times \delta Y = - \tilde{F}, \quad (19)
\]

where \( \tilde{Y} \) denotes the approximate unknown vector, \( \delta Y \) is the correction vector, \( \tilde{F} = (\tilde{F}_{1}, \tilde{F}_{2}, ..., \tilde{F}_{N})^{T} \) is the RHS vector of the system calculated by \( \tilde{Y} \) and \( [\partial F / \partial Y] \) is the \( N \times N \) Jacobi matrix with the Newton method derivatives. In the case of normal convergence, both \( \delta Y \) and \( \tilde{F} \) approach to zero.

The Jacobi matrix, obtained by differentiation of Eqs. (14)–(17), has a tridiagonal structure with filled first row and column and zero element in the upper left corner. Four terms in the last line and the last column are not zero in the case of cyclic boundary conditions for Eq. (17):
where \( c \equiv \hbar^2/(2m\Delta x^2) \) and \( a_i \equiv E - V_i - 2c \).

In the case of zero boundary conditions of Eq. (16), in the last line of the matrix (20) there remains only one non-zero element \( a_N = 1 \). A more radical way is to solve the problem with the \((N-1)\times(N-1)\) matrix without the last row and column for the respectively reduced unknown vector, since \( \Psi_N = 0 \) is fixed and must not be considered as an unknown variable.

To solve the linear system (19) with the tridiagonal bordered matrix (20), a special effective Gaussian elimination algorithm ([13], p 90) was applied.

Using approximate form for the wave function and the value of the energy level, the EWC method described here finds iteratively the exact value for energy and wave function for any specified potential. The approximate initial solution may be obtained from theoretical estimations or with any other method. The only drawback of the EWC method is that in the case of a poor initial approximation when, e.g., a wrong number of halfwaves of the wave function within the QW is determined, the convergence process may give a solution for a different eigenstate instead of the one which was wanted. The method was tested in the case of MQW structures with hundreds of abrupt potential barriers [9]. A reliable scan over the whole range of eigenenergies was obtained if rough approximations were precalculated by the shooting method.

Numerical tests showed that in the case of poor initial solutions the EWC method usually needed 5–7 iterations to converge. Divergence was never observed. In a combined use together with the shooting method, the typical number of Newton iterations decreased to 3–4 to reach convergence with practically zero error. Figure 2 illustrates very high quadratic convergence speed of the EWC method.
3. RESULTS OF THE EFFICIENCY TEST

3.1. Comparison of the accuracy of the FGH and EWC methods

Prior to comparing the efficiency on the basis of consumption of the computer time, an estimation of equivalent spatial grid sizes should be performed. Some preliminary calculations in the case of a nearly sine wave function showed that minimal rough accuracy may be achieved with the FGH method if the number of grid nodes per wave of $\Psi$ is 2–3. For EWC method the corresponding number was 4–5. A more accurate comparison was performed for a triple QW task as described in Fig. 3.

The structure in Fig. 3 has three quantum wells of 9 Å width and 10 eV depth, separated by 1 Å barriers. The size of the outer barriers on both sides is 9 Å. The structure has 15 energy levels below 10 eV. The separation of energy levels has peculiarities and the wave functions contain intervals with different spatial frequencies. In tests, the free electron rest mass was used. To achieve maximum compatibility between the FGH and EWC methods, which use slightly different localization of spatial grid nodes, special care was taken. Additionally, for the border nodes in the FGH method a high potential energy

![Fig. 2. Typical convergence characteristics of the iteration process of the EWC method. Figure shows in logarithmic scale the decrease of the bound state energy increment $\delta E$ (eV), the decrease of the maximum relative wave function increment and the approach of the particle finding probability to its theoretical limit of 1.](image-url)
value was assigned to model the zero boundary conditions of the EWC method. In this comparison the accuracy of both methods was evaluated by a comparison of the accuracy of the energy eigenvalues calculations. Since the computer time consumption of the FGH method became very high for greater grid point numbers, exact reference numbers were obtained with the EWC method for the grid with \( N = 30 \, 000 \). The accuracy criterion for finishing the Newton iterations in the EWC method was practically set to zero \((\delta E \leq 10^{-3} \, \text{eV})\). The results are presented in Fig. 4.

As Fig. 4 shows, for both methods the error of energy level decreases with the grid step as \( -\Delta x^2 \). However, to achieve a comparable accuracy, the EWC method needs approximately three times more grid nodes. For example, a relatively good accuracy of 0.1 meV needs approximately 1500 nodes in the case of the FGH method and 4500 nodes in the case of the EWC method.
3.2. Comparison of computation times of the FGH and EWC methods

The computer times for both approaches are compared in Fig. 5. The calculations were performed for a triple QW structure with 15 energy levels according to Fig. 3. As the approximate wave functions and energy eigenvalues for the EWC method were calculated by shooting method, the corresponding results are marked as ShM+EWC in Fig. 5. In the case of the FGH method, the computer time practically depends only on the number of grid points and not on the form of the potential. In the case of the EWC method, the amount of computer time is also proportional to the number of energy levels. Numerical experiments were performed on a desktop PC with a 3 GHz Pentium-4 processor on Windows XP platform using GNU-Fortran-77 programming language.

Figure 5 shows the $t_{\text{comp}} \sim N^3$ dependence of the FGH method. That limits the practical use of this method to the grid sizes from 2000 to 3000. It is interesting that the reduced version of the FGH, which does not calculate the wave functions, does not significantly reduce the computational time. In contrast to the FGH, in the case of the EWC method $t_{\text{comp}}$ depends linearly on $N$. Although the EWC method needs roughly three times more grid nodes for the same accuracy,
the comparison proves clearly that the ShM+EWC approach is more than three orders of magnitudes more effective than the FGH method. Computer times of subsecond range show that the ECW method may be easily applied to very complex MQW problems that demand $10^5$ and more grid nodes.

Some tests were also performed with the shooting method on its own. However, the exact results of a comparison between the ShM and the combined ShM+EWC method depend on the specific finishing criteria for the coordinate and energy iteration processes in the ShM. The EWC method is rather insensitive in the sense that its boundary conditions are clearly fixed and the energy convergence speed is very high (Fig. 2). Numerical experiments showed that by careful selection of the accuracy criteria and using the benefit of a symmetric structure, it was possible to obtain 30–40% shorter computer times with the pure ShM. However, in the general case of a non-symmetrical structure and guaranteed high accuracy the combined ShM+EWC approach was approximately 1.5–2 times more effective than the use of the ShM only.
4. DISCUSSION AND CONCLUSIONS

We have discussed and tested qualitatively some practical approaches to the solution of the time-independent 1D Schrödinger equation without any restrictions on the potential energy distribution. The comparison was centred around two effective methods: the Fourier grid Hamiltonian method known from the field of physical chemistry and the shooting method often applied to semiconductor quantum well calculations. We established cubic computer time dependence on the number of grid points in the FGH method and concluded that it is very difficult to use this method for complex tasks which need more than 2000 grid points.

We also critically analysed the drawbacks of extremely simple trial-and-correction type shooting methods and offered a more general and reliable coupled energy and wave function method (EWC) with a Newton iteration scheme and an internal linear task with a tridiagonal bordered matrix. We formulated this EWC method for two types of boundary conditions: zero wave function (hard wall) or cyclic. A survey of the literature showed that the EWC method was actually an extended version of a relaxational approach \[^4\] published in 2001.

For versatile multiquantum well problems we developed an effective and reliable combined approach (ShM+EWC) where at first the shooting method is used for a rough estimation of the energy eigenvalues and approximate wave functions. Secondly, the fast-converging EWC method is applied for reliable calculation of more exact results. Detailed investigation of the grid error and computer time on the basis of a triple quantum well task for both the FGH and ShM+EWC approaches was performed. The results show that although the ShM+EWC method needs approximately three times more grid nodes than the FGH method, it is still several orders of magnitudes more effective than the FGH method. On modern computers, MQW tasks with grid point numbers over \(10^7\) may be easily solved with the EWC method.

ACKNOWLEDGEMENTS

This work has been partly supported by the Estonian Science Foundation (grant No. 5911). Paul Harrison would like to thank the Faculty of Engineering of the University of Leeds for financial support. The authors are grateful to Dr. Zoran Ikonić from University of Leeds for valuable comments on the algorithms for solving the Schrödinger equation.

REFERENCES

10. EISPACK package of Fortran subroutines that compute the eigenvalues and eigenvectors of matrices (developed in 1972–73), http://www.netlib.org/eispack/

Schrödingeri võrrandi lahendusmeetodite võrdlus mitmik-kvantaukudega heterostrukturiude jaoks

Andres Udal, Reeno Reeder, Enn Velmr ja Paul Harrison

On vörrelud otseseid numbriilisi lahendusmeetodeid ajast sõltumatu ühemõõtmelise Schrödingeri võrrandi lahendamiseks. Mitmik-kvantaukudega (MQW) poolijuht-heterostrukturiüülikidest arvutused nõuavad meetodeid, mille puhul arvutusaj t_{comp} sõltub ruumivörgu sammude arvust N lineaarselt. Tuntud ja väga efektiivseks peetav Fourier Grid Hamiltonian (FGH) meetod (Fourier’ teisenduse ja ruumivörgu alusel moodustatud hamiltoniaani analüüsiv meetod) omab aga kuup-sõltuvust t_{comp} \sim N^{1}, mistõttu selle meetodi rakendusala on piiratud probleemidega, kus N \leq 1000 on piisav. Lihtsaim otsene lahendusmeetod on nn tulistamismeetod (ShM), mis põhineb katseselast astumisel üle ruumikoordinaadi ja energiaviärtustest. Tulistamismeetod omab vajalikuks lineaarsest sõltuvust t_{comp} \sim N ja rahuldavat energiaviärtust koondumiskiirust, kuid ebaselgelt määratletud piiritingimused teevad meetodi kasutamise ebamugavaks. Artiklis on esitatud energianivoode ja lainefunktioonide kooslahendamide meetod (EWC), mis on töökindel ja efektiivne ning omab lineaarsest sõltuvust t_{comp} \sim N. Meetod põhineb mittelineaarsete võrrandi-süsteemide lahendamiseks sobival Newtoni iteratsioonimeetodil, kusjuures sise-

260
mise lineaarse ülesandena lahendatakse kolmediagonaalse ääristatud maatriksiga süsteem. Esitatud meetod on rakendatav suvalise potentsiaalse energia jaotusega ülesannetele keerukusega \( N = 10^5 \) ja üle selle nii nulliliste kui ka tsükliliste piirtingimuste puhul. Praktiliste MQW-ülesannete jaoks on realiseeritud võimalus kasutada meetodeid kombineeritult, mille puhul arvutatakse ShM-i abil energiate ja lainefunktsioonide ligikaudsed algläheid väga kiirelt koonduvale EWC-mетодile. Kolmik-kvantu güntel formuleeritud testülesande lahendamise teel on vörreldud vaadeldud meetodite ruumilist täpsust ja arvutiaega. Tulemused näitavad, et kombineeritud meetod ShM + EWC on FGH-meetodist mitu suurusjärku efektiivsem.
Laterally pumped GaAs/AlGaAs quantum wells as sources of broadband terahertz radiation

Reeno Reeder
Department of Electronics, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

Zoran Ikonić and Paul Harrison
Institute of Microwaves and Photonics, School of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, United Kingdom

Andres Udal and Enn Velmre
Department of Electronics, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

(Received 28 February 2007; accepted 23 July 2007; published online 11 October 2007)

In this work we consider lateral current pumped GaAs/AlGaAs quantum wells as sources of incoherent terahertz radiation. The lateral field heats the electrons in a two-dimensional quantum layer and increases the population of higher subbands, hence also increasing the radiation power generated in spontaneous intersubband emission processes. Digitally graded quasi-parabolic and simple square quantum wells are considered, and the advantages of both types are discussed. Calculations at lattice temperatures of 77 K and 300 K, for electric fields up to 10 kV/cm, show that the optical output power of ~100–200 W/m² may be achieved for the 7 THz source. The main peak of the spectrum, at 7 THz, of the quasi-parabolic quantum well exceeds the black body radiation at 300 K by approximately a factor of two and by two orders of magnitude at 77 K.


I. INTRODUCTION

The development of sources of terahertz radiation has become a hot topic in the last decade because of numerous prospective applications. The successful realization of terahertz quantum cascade lasers has made a huge impact on Terahertz technology. These coherent sources are able to give good levels of output power but still require low temperatures for their operation [below 164 K (Ref. 1)]. However, not all possible applications of terahertz radiation really require coherent and highly monochromatic sources. In fact, it is sometimes advantageous to have a reasonably broadband terahertz source that would preferably operate at room temperature. In this work we consider the possibility of using a relatively simple quantum well, pumped by a lateral current, as a source of incoherent terahertz radiation. This would be generated in spontaneous radiative transitions between size-quantized states and requires just electron excitation into higher states (subbands) and not any population inversion. The bandwidth of such sources is limited from below by the spontaneous transition line width, i.e., is intermediate between those obtained from lasers and from thermal sources. However, a more broadband output can be achieved by appropriately engineering the subband structure. Their emitting area can be quite large, generally limited only by the wafer size. In structures based on conventional III/V materials like GaAs/AlGaAs, the generated radiation is intrinsically polarized perpendicularly to the well layer, hence propagating in this plane, but surface-normal emission is possible by using grating-type outcouplers.

II. THEORETICAL CONSIDERATIONS

Electrons in excited subbands of a quantum well relax into lower ones mostly by scattering processes (with phonons, interface roughness, etc.) and also, though with a small efficiency, by spontaneous emission of photons. The emitted radiation is line-like, its bandwidth determined by the line broadening. Some amount of radiation is emitted even under equilibrium due to a finite population of excited subbands at any finite temperature, and the power spectrum is also line-like, because a quantum well is not a black body. Driving the system off equilibrium, e.g., by applying a lateral electric field (Fig. 1), will generally increase the population of excited subbands and hence, the generated optical power. To get a comparatively narrowband (i.e., spontaneous-emission-width-limited) emitter of this type, one can pursue one of the following two approaches. One is to use a quantum well with a number of equispaced subbands, e.g., a parabolic-like quantum well, and rely on the fact that the only strong optical transitions are those between adjacent states. In an ideal parabolic potential only these transitions

![FIG. 1. Laterally pumped quantum well emitter. The doping is localized in a 2D quantum layer and in the regions under the contacts.](image-url)
are indeed allowed (have nonzero dipole matrix elements), but the situation is similar for many other conventional potentials. Any strong optical transition in such a system would thus contribute to the same emission profile, which would thus be independent of how strongly the system is pumped, i.e., what number of higher subbands are significantly populated. The other approach would be to use a quantum well with the lowest two states spaced by the required amount, while the third state would be much higher up and therefore unlikely to acquire a significant electron population, the situation which effectively guarantees a narrowband emission spectrum.

In either case, there is clearly a tradeoff between the pumping strength (and hence the emitted optical power) and the degree of monochromaticity of the emitted radiation, the latter being compromised by electron excitation into high subbands which cease to be equispaced (in the case of quasi-parabolic wells), or into the third or higher subbands (in case of a square well). To address this question, we performed modeling of subband population dynamics in appropriate quantum wells subject to an in-plane electric field. The relevant physics is shown in Fig. 2. Electrons obtain kinetic energy from the in-plane electric field, their in-plane distribution being shifted upward along the subband. The change of the distribution modifies the electron scattering rates. In particular, for sufficiently large fields the “upward” (i.e., subband-elevating) scattering rates will clearly increase, because a large fraction of electrons in a lower subband will have a large enough kinetic (and hence also the total) energy to be above a higher subband near its zone center. Electrons thus climb up the subband ladder, which results in an increased population of higher subbands compared to the equilibrium case. Electrons in higher subbands relax into lower ones not only by scattering but also by spontaneous photon emission. The generated optical power depends only on subband populations, not on the in-plane electron distribution, because optical transitions are vertical (electron wave vector conserving) and, within the parabolic in-plane dispersion model, pairs of subbands are essentially equispaced for all wave vectors. It is worth noting that considerations of hot electrons in a square well as a source of terahertz radiation have been presented in Ref. 3 but within a simple model that assumed equilibrium electron distribution throughout the system, with their temperature taken as an input parameter.

The electron distribution due to an in-plane electric field in a multisubband system may be found by solving the Boltzmann equation with all intra- and intersubband scattering processes taken into account. In this work, however, we use a simplified and much faster approach, taking the in-plane electron distribution (over the wave vectors of a subband) to be described by the shifted Maxwell–Boltzmann (MB) distribution and then explicitly handling only the intersubband scattering processes. Detailed calculations of the electron transport in bulk semiconductors show that the shifted MB distribution (see Fig. 3), with appropriate field-dependent temperature and drift velocity in the direction of the field, can be a very good approximation to the actual calculated distribution. It has also been previously used, in form of a shifted Fermi–Dirac (FD) distribution, in single-subband electron-transport calculations in quantum wells. In this work we apply such a model to systems with a number of size-quantized subbands and assume that each of them will have the same form of shifted-MB distribution over the in-plane wave vectors, which is then used to evaluate the distribution-averaged intersubband scattering rates, Eq. (9.153) from Ref. 6.

It should be noted that the validity of this approximation relies on having strong electron-electron intrasubband scattering, because this is the major process which brings about the electron thermalization (by which we mean that the distribution acquires the MB- or FD-like form, with its temperature generally different from that of the lattice). Under the operating conditions typical for the devices we consider ($T = 77–300$ K, area doping $\sim 10^{12}$ cm$^{-2}$), estimates of the electron-electron intrasubband scattering time constant put it into the deep subpicosecond region. This is much faster than any other intra- or intersubband scattering process involved in this work (slightly subpicosecond timescale, at best), and therefore using the shifted MB distribution approximation can be justified. On the other hand, it may not be necessary to explicitly include the intersubband electron-electron scattering in the rate equations. The rate of this process decreases with intersubband spacin,
indeed accounted only for polar LO phonon and acoustic phonon intersubband scattering, the former being the major intersubband scattering process. Since the calculation of electron-electron scattering in a multisubband system is far slower than for phonon-electron scattering, this makes large savings of computation time. At any reasonable value of the field (high enough to give technically significant effects in the system under consideration) the electron temperature is large enough that the MB and FD distributions (both shifted, in this instance) are essentially indistinguishable, and we use the latter form just for convenience.

The electron energy in the $n$th subband is $E_n = E_{n0} + \hbar^2 k^2 / 2m^*$, where $E_{n0}$ is the size-quantized subband minimum energy, and the kinetic component depends on the two-dimensional (2D) in-plane wave vector $k$. According to the shifted-FD distribution model, the electron distribution over the in-plane $k$ states depends not simply on their energy but is rather described as

$$f^{FD}(k) = \left[ 1 + \exp \left( \frac{E_{n0} + \hbar^2 [(k_x - k_{0x}(F,T_{latt}))^2 + k_y^2] - E_{Fx}}{2m^* k_BT_{latt}(F,T_{latt})} \right) \right]^{-1},$$

where $k=(k_x,k_y)$ is the in-plane wave vector, $E_{Fx}$ is the quasi-Fermi level of the $n$th subband, while the drift wave vector $k_0$ and electron temperature $T_{el}$ both depend on the field $F$ (here taken in the $x$ direction) and the lattice temperature $T_{latt}$. For the numerical calculations the relations $k_{0x}(F,T_{latt})$ and $T_{el}(F,T_{latt})$ may be obtained using data from the literature (e.g., Refs. 2 and 8–10), partly relying on the fact that the energy loss rate of the electron gas versus electron temperature dependence is almost the same for 2D and 3D systems.11 Figure 4 shows the electron drift velocity $v_d = \hbar k_0 / m^*$ and temperature dependencies on the field, as used in this work. These were assembled using the theoretically calculated and experimental data from several sources.2,8–10

At equilibrium ($F=0$) the (quasi-)Fermi level is the same for all subbands, but for nonzero electric fields the subband populations $n_i$ change, and the Fermi level for each subband is found from

$$n_i = \frac{2}{4\pi^2} \int_{k_{ix},k_{iy}} f^{FD}(k) dk_x dk_y,$$

The rate equations describing the subband population dynamics in the steady state read

$$\frac{dN_f}{dt} = \sum_{i=1}^{N} \frac{1}{\tau_{if}} [n_i - n_f] \sum_{j=1}^{N} \frac{1}{\tau_{ji}} = 0,$$

where $\tau_{ij}$ is the total averaged scattering rate from the $i$th to the $f$th subband, due to all scattering processes accounted for, and $N$ the number of subbands included in the model. There are $N$ such equations, making a homogeneous system. Any one of these is linearly dependent on all the others and is replaced by the particle conservation law (i.e., the global charge neutrality), $n_1 + n_2 + \ldots + n_N = N_{el}$, where $N_{el}$ is the overall electron concentration per unit area, set by the structure doping. This makes a standard inhomogeneous linear system of equations, to be solved.

In deciding on the number of subbands to be included, one normally takes all those low-lying subbands that are expected to host a non-negligible electron population. Typical quantum well structures have a limited number of bound states (subbands), of the order of 10 or so, and including all of them present no difficulties on the computational side. However, if the actual calculation shows that even the highest bound subband becomes significantly populated at some value of the in-plane field, this is a signal that population of the above-the-barrier (continuum) subbands has begun. The model has then to be expanded to include a sufficiently large range of energies in the continuum. Since we have encountered such an operating regime, the model includes a range of continuum states, along with all the bound subbands, in the description. The continuum is described by embedding the real quantum well in a box much wider than the well width, which produces a dense spectrum of quasi-continuum states. Including any reasonable range of continuum states grossly increases the total number of states that are explicitly handled, i.e., the order $N$ of the system [Eq. (3)], but this has to be done for sufficiently large fields.

The subband populations are thus found from the rate equations, using the distribution-averaged intersubband scattering rates, which themselves depend on the Fermi levels.6 With this cross-dependence, it is clearly necessary to use an iterative (self-consistent) solution of the whole system, starting with arbitrary initial values, until the required precision is achieved.

With the subband populations found, the total emitted (more precisely, internally generated) optical power, due to spontaneous intersubband transitions, is calculated from

![FIG. 4. Electron drift velocity and temperature dependencies on the lateral electric field used in this work, assembled from various theoretical and experimental data (Ref. 2 and 8–10).](http://jap.aip.org/jap/copyright.jsp)
\[ P_{\text{total}} = \sum_{i>j}^{N} \frac{n_i}{\tau_{ij}^{\text{rad}}} \delta_3 E_i - E_j, \]

where \( \tau_{ij}^{\text{rad}} \) and \( \hbar \omega_{ij} = E_i - E_j \) are the radiative lifetime and energy spacing between \( i \)th and \( j \)th subband, respectively. The radiative lifetime is given by

\[ \frac{1}{\tau_{ij}^{\text{rad}}} = \frac{e^2 \bar{n} (E_i - E_j)^3 d_{ij}^2}{3 \pi \varepsilon_0 c \hbar^2}, \]

where \( \bar{n} \) is the refractive index, and \( d_{ij} \) is the optical dipole matrix element, \( d_{ij} = \int \Psi_i^*(z) \Psi_j(z) dz \). The Lorentzian-type power spectrum is calculated from

\[ P(\hbar \omega) = \sum_{i>j}^{N} \frac{n_i}{\tau_{ij}^{\text{rad}}} \frac{\Gamma}{\pi (\hbar \omega - \hbar \omega_{ij})^2 + \Gamma^2}, \]

where \( \Gamma \) is the linewidth (half width at half maximum) of the intersubband transitions, here set to 5 meV as a typical value.

### III. Numerical Results and Discussion

In this work, the emission characteristics of the digitally graded, quasiparabolic and square wells were investigated. Continuous composition-graded wells which would provide a truncated parabolic potential, based, e.g., on Al\(_x\)Ga\(_{1-x}\)As alloys, are quite difficult to grow. Furthermore, a strictly parabolic composition grading does not lead to equispaced states because of the effective mass variation, though this can be corrected by modifying the composition profile.\(^{12,13}\) For these reasons we have tailored a more readily realized, digitally graded quantum well, with the layer widths carefully adjusted to give 12 equispaced states to high accuracy, with the state spacing equal to approximately 29 meV, corresponding to a 7 THz radiative transition frequency.\(^3\) The structure is shown in Fig. 5. The layer widths are integer multiples of crystalline monolayer (0.283 nm), which should minimize interface roughness.

Calculations were performed for the lattice temperatures of 300 and 77 K. Doping of \( 10^{12} \text{ cm}^{-2} \), only within the quantum well layer structure, was assumed. Using just the 12 equispaced bound states proved insufficient; hence, an additional 24 states of the quasicontinuum, covering the energy range of 0.110 eV above the barrier top, were also included in the model, because a sizeable fraction of electrons gets excited into these continuum states at larger fields. Certainly, the equispaced-states property does not apply to the continuum. The confinement box size for the quasicontinuum states was selected as 185 nm, approximately 3 times greater than the width of the QW.

In Fig. 6 we give the total generated power dependence on the lateral electric field. This starts to increase significantly only beyond 1 kV/cm (2 kV/cm for 77 K), and enters a saturation for biases in excess of 8 kV/cm. The power range obtained, \( 100-200 \text{ W/m}^2 \), is limited by the general cubic dependence on the energy spacing [as follows from Eqs. (4) and (5)], together with the fact that \( d_{ij}^2 \sim 1/\hbar \omega_{ij} \). The power spectrum also evolves with the field and becomes progressively less narrow, as shown in Fig. 7 for the quasiparabolic structure. For fields up to about 4-5 kV/cm the spectrum is rather “monochromatic,” with a single peak around the required frequency of 7 THz, as expected from a light-emitting diode (LED) device. For higher fields, however, the side-peaks at 21 THz and 35 THz start to contribute to the output power. One can easily estimate, assuming the unity value for the surface emissivity, that the total power within the 7 THz peak at 10 kV/cm exceeds the black body radiation in the same spectral range by approximately a factor of 2, even at room temperature (300 K), and grossly exceeds it, by 2 orders of magnitude, at 77 K.

**FIG. 5.** Digitally graded quasiparabolic quantum well for 7 THz emission (Ref. 13). The equivalent ideal parabolic well potential is also shown. The layer widths in crystalline monolayer units, starting from the middle one (because the structure is symmetric), are as follows: 12, 1, 11, 1, 8, 1, 1, 2, 1, 2, 1, 4, 2, 1, 7, 1, 2, 4, 1, 4, 3, 1, 2, 5, 7, 3, 1, bulk (GaAs - boldface, Al\(_{0.42}\)Ga\(_{0.58}\)As - regular font).

**FIG. 6.** Total generated optical power dependence on the lateral electric field, calculated for the digitally graded, quasiparabolic well, and for the simple square well for lattice temperatures 77 K and 300 K.

**FIG. 7.** (Color online) Generated power spectrum for the digitally graded quantum well. The 11 displayed spectral profiles are for electric fields of 0, 1, 2, ... \( \ldots, 10 \text{ kV/cm} \).
The appearance of higher frequencies is clearly related to the increasing electron population in higher, continuum subbands, as shown in Fig. 8, which becomes quite prominent for fields above 3 kV/cm. It should be noted, however, that it may not be just electron escape to continuum which contributes to spectrum broadening. It is only the ideal parabolic potential, and not just any equispaced-states potential, which has the property that no other transitions except those between adjacent states are allowed. Indeed, starting with the parabolic potential one can generate families of asymmetric equispaced-states potentials which have quite significant transition matrix elements between more remote states. However, in the digitally graded well considered here the dipole matrix elements between subbands with the quantization index changing by three are between 1 and 2 orders of magnitude smaller than those between the adjacent subbands, and others are much smaller; hence, this path for the power spectrum broadening is still of minor importance: the major part of broadening comes from the transitions from continuum states.

Another set of calculations was performed for a 20 nm wide GaAs square well embedded in a Al$_{0.3}$Ga$_{0.7}$As barrier. The well width was chosen so that the energy spacing of the first two subbands is about the same as in the quasiparabolic well, i.e., approximately 30 meV. This structure has five bound subbands, at 10.1, 40.4, 90.5, 159, and 238 meV. It was embedded in a 220 nm wide outer box, and 31 quasi-continuum states were also included in the model. The calculated emission characteristics are shown in Figs. 6 and 9, and the relevant subband populations in Fig. 10. Clearly, the total power generated in this structure exceeds that of the quasiparabolic well, but the spectrum is much broader.

**IV. CONCLUSION**

Laterally pumped GaAs/AlGaAs quantum wells were considered as potential broadband sources of incoherent terahertz radiation, generated on spontaneous intersubband optical transitions. The in-plane field heats the electrons and hence increases the population of higher subbands, and therefore also the output power. Two types of quantum wells were considered, a digitally graded, quasiparabolic well and a simple square well, and their performance was compared. Despite the higher overall emitted power from the square well system, the emission spectrum of the quasiparabolic well is generally narrower, and up to the moderate fields remains rather monochromatic, as expected from spontaneous emission-based sources like LED devices. The total optical output power at 7 THz is limited due to the cubic dependence of the spontaneous emission on the subband energy spacing. However, the spectral power at 7 THz still exceeds, by approximately a factor of two, the ideal black body radiation at 300 K. The analysis presented predicts that this device, in contrast to quantum cascade lasers, may operate at room temperature, or even at 400 K. A higher optical output may be achieved either by increasing the electron density or by using a stack of multilayer structures. Furthermore, by appropriate modifications of the digitally graded quantum well, it appears possible to realize limited-bandwidth incoherent terahertz sources, the spectral properties of which are controllable by the built-in GaAs/AlGaAs composition distribution, and also dynamically (to an extent) by the lateral field.

**ACKNOWLEDGMENTS**

The authors are grateful to N. Vukmirović for helpful discussions. This work has been partly funded by Estonian.
Science Foundation grants 5911 and 6914, the Estonian Archimedes Foundation, and the Engineering and Physical Sciences Research Consul (U.K.) THz grant EP/C002881/1.

6P. Harrison, Quantum Wells, Wires and Dots (John Wiley & Sons, Chichester, 2005).
Article IV

Advances in Terahertz Technology with Emphasis on Quantum Cascade Lasers

R. Reeder, E. Velmre
Department of Electronics, Tallinn University of Technology,
Ehitajate tee 5, 19086 Tallinn, Estonia, phones: +3725278075, +3726202162; e-mails: reeno.reeder@gmail.com, evelmre@ttu.ee

A. Udal
Department of Computer Control, Tallinn University of Technology,
Ehitajate tee 5, 19086 Tallinn, Estonia, phone: +3726202110; e-mail: audal@va.ttu.ee

Introduction

After several rises and falls of enthusiasm in terahertz domain during many decades [1], it seems that at last the T-rays research and technology has found the real boost during last 10 years. At that an important role have had the long time waited solid state source - the Quantum Cascade Laser (QCL) which idea was proposed already in 1971 by R. Kazarinov and R. Suris but the real working devices were realized in mid-infrared region region (3-5 μm) only in 1994 [2] and in THz-region in 2002 [3]. The increase of relevant research during last 10 years has been impressive, see Fig.1.

Historically the THz range was often called as “terahertz gap” because of the lack of sources in this part of spectrum. Solid-state electronics capability was limited for a long time at terahertz frequencies, for example, in 1922 the unexplored region was between the wavelengths of 0.2 and 2 mm [9]. The early emitters in this range were either very inefficient or their development was complicated [1], leaving the beginning of the race in this area to the end of 20th century. In 2005 the clear gap in power still existed in the field of the solid-state THz sources [1], see Fig.2.

Terahertz frequency area promises many new applications on spectroscopy, terahertz imaging and microscopy, genetic sensing, detection of biological and explosive hazards, astronomical telescopes [4], high-speed wireless communications [5, 6], and even real-time terahertz imaging at video rate [7, 8].
Advances in Terahertz Sources

The variety of terahertz sources is quite large by today. Narrow-band THz radiation can be produced by free-electron lasers and fast diodes. Broadband terahertz radiation can be produced by thermal sources, laser-driven sources, and by short electron bunches in accelerators [14]. Depending on the type of application, either narrow or broadband sources can be used. Most of the possible applications expect the emitter to have high power and small dimensions, and to be cheap and portable. Today, there are several techniques that provide high power, but the equipment for that is often not very compact, being rather built in to a laboratory, than portable.

The first source best satisfying the need for high power is Free Electron Laser (FEL). It is a monochromatic tunable oscillator having good power but unfortunately it is unportable. There are reports that say about 400 watts of average and 600 kW of pulse power in the range of 1.3 – 7.5 THz using FEL [15] that are remarkably good result. Another report tells about a 1-watt average result in the range of 0.5 – 1.5 THz [16].

Another way to produce high power broadband THz radiation is using subpicosecond electron bunches in an accelerator. Using this technology 20 watts of power have been measured in the terahertz region [14].

An also good technology yielding relatively good power in THz band is gyrotron, which use vacuum tubes that emit terahertz beams by bunching electrons with cyclotron motion in a strong magnetic field [17]. In the experiments by Idehara et al the frequency of the gyrotron is tunable between 120 GHz and 1080 GHz having magnetic field of 20 T at the highest frequency, giving output power from several tens of watts to several hundreds of watts [18,19]. The equipment of these experiments is very large, occupying even several floors, but making the device compact is under progress.

The next way to generate the radiation is using $p$-Germanium laser which is based on population inversion between light and heavy holes. Up to 100 mW of power in the range of 170 – 200 μm (1.7 – 1.5 THz) have been reported [20]. A bad side of this source is the requirement for low temperatures near 15 K [21].

Quite an old method to produce terahertz emission is using CO$_2$ gas laser, either based on molecular pumping [22] or mixing of different frequencies [23]. Molecular pumping have a relatively broad spectrum, ranging from 0.580 to 4.25 THz, giving 20 – 30 mW output power [22].

Another method is to use tunnel injection transit time diodes, which are so called TUNNETT diodes [24]. Reports are showing up to 140 μW power in 0.355 THz frequency [25].

Backward-wave oscillators that use biplanar interdigital slow wave circuits are reported to give 23.8 mW near 0.650 THz [26].

Up to 1.1 mW power in the range of 3.2 – 4.8 μm have been obtained [27] using the difference frequency generation. This is a principal method to get higher frequency $\omega_3$ by summing two lower frequencies $\omega_1$ and $\omega_2$ [28].

Optical rectification which is based on mixing two frequencies, was first showed in 1962 [29]. This method is based on the inverse process of the electro-optic effect.

Terahertz radiation from nano-accordion motion of standing waves in an impedance-mismatched layer sandwiched by GaN-based piezoelectric heterostructures have been investigated recently [30]. This may lead us to a new way to produce terahertz radiation.

By combining monolithic microwave integrated circuit amplifier frequency tripler chips, output power > 1 mW have been demonstrated at around 0.9 THz [31].

Simulations have been done by the authors of the present work too using spontaneous radiation from laterally pumped quasiparabolic GaAs/AlGaAs quantum wells [32]. The results show at 7 THz ca 100-200 W/m$^2$ output power couple times over blackbody radiation up to temperature 400K.

Advances in Development of Quantum Cascade Lasers

A small hop can be seen near 1994 when the MIR Quantum Cascade Laser was demonstrated by Faist et al [2]. Shift to the THz domain in 2002 [3] initiated also the terahertz era in semiconductor emitters. After that a bigger hop in general was near 2004 caused by commercial sources coming to the market. This boosted up the publishing of results on application related experiments.

Currently, Quantum Cascade Lasers are the only semiconductor devices operating from the mid-infrared region to the THz range of frequencies [33].

The main problems with QCL-s are their low working temperature and small power. The main emphasis is on developing these properties – to increase working temperature, where they are still effective, and to increase the power in general.

During last couple years, the THz QCL-s have been quickly approaching the higher temperatures and longer
wavelengths by applying different sophisticated design ideas and the strong external magnetic field. By 2008 the maximum operating temperature without external magnetic field was increased to 178 K in pulsed mode and 117 K in continuous wave (CW) mode at the lowest frequency of 1.2 THz [34].

In 2009 demonstrated Wade et al. (5 leading research centers of USA) that with resonant phonon desing and with strong magnetic field (over 16 T) may be achieved QCL work in 1 THz regime at temperatures up to 215 K, and 3 THz regime at up to 225 K [35]. In the resonant-phonon design scheme, the population inversion is ensured by selectively injecting electrons through resonant tunnelling into the upper state of the laser transition [35].

The room temperature lasers exist in MIR range [36], e.g. InGaAs-AlAsSb QCL 3-4 μm, several mW at 270 K [36]. Using metal grating distributed feedback, continuous wave room temperature operation at wavelengths 4.5-7.5 μm and output power 20 mW is achieved [37].

The general trends of THz QCL development are illustrated in Fig.4 below. At longer wavelengths the design problems become especially complicated if the temperature energy exceeds the energy quantum:

\[ k_B T \geq h \nu = h \frac{\nu}{c} . \]  

(1)

Fig. 4. The achieved results and general trends of development of THz quantum cascade lasers

The output frequency of QCL-s can be tuned either by bias voltage [38] or by external cavity properties [39]. Recent achievements include also Quantum Cascade Lasers with Ultra-Strong Coupling Injection with peak power of 8W (2W) at 80K (300K) [40].

Conclusion

Since the year 2002 the development of THz Quantum Cascade Lasers has become an extremely dynamic research field. The present paper discusses the recent advances in this area.

Acknowledgement

The authors wish to thank the Estonian Ministry of Education and Research (the target oriented project SF0142737s06), the Estonian Science Foundation (6914), and the Foundation Archimedes through the Centre of Excellence CEBE (TK05U01) for supporting the presented research work.

References


Received 2010 02 15


Starting from the year 2002 when the development of quantum cascade lasers shifted from mid-infrared to the terahertz domain, the remarkable rise in the terahertz research has followed. The paper discusses the recent achievements in development of the terahertz radiation sources with emphasis on quantum cascade lasers. Moving towards the higher working temperatures, longer wavelengths and higher output power is discussed. Ill. 4, bibl. 40 (in English; abstracts in English, Russian and Lithuanian).
Article V

Development of quantum cascade laser simulation software

A. Udal\textsuperscript{1}, R. Reeder\textsuperscript{2}, Z. Ikonie\textsuperscript{3}, P. Harrison\textsuperscript{4}, E. Velme\textsuperscript{5}

\textsuperscript{1}Department of Computer Control, TUT, Ehitajate tee 5, 19086 Tallinn, Estonia. E-mail: andres.udal@ttu.ee
\textsuperscript{2}Department of Electronics, TUT, Ehitajate tee 5, 19086 Tallinn, Estonia. E-mail: reeno.reeder@gmail.com
\textsuperscript{3}School of Electronic & Electrical Engineering, University of Leeds, LS2 9JT, U. K., E-mail: Z.Ikonie@leeds.ac.uk
\textsuperscript{4}School of Electronic & Electrical Engineering, University of Leeds, LS2 9JT, U. K., E-mail: P.Harrison@leeds.ac.uk
\textsuperscript{5}Department of Electronics, TUT, Ehitajate tee 5, 19086 Tallinn, Estonia, E-mail: evelme@ttu.ee

ABSTRACT: Quantum cascade lasers are the sophisticated modern semiconductor devices of quantum electronics needed for coherent radiation generation in mid infrared and terahertz regions. The development of respective simulation software needs special efforts of quantum and software engineering to make the simulation tools usable not only on powerful computer clusters but also on desktop computers. The present short paper and the corresponding conference presentation discuss the development of desktop simulator QCLSIM1 for mid-infrared AlGaAs/GaAs quantum cascade laser structures.

1 Introduction

The first semiconductor lasers were realized in autumn 1962 nearly simultaneously in USA by the four research groups [1]. The coherent laser radiation was obtained in pulsed mode at low temperatures from the basic simple the pn-junction diode structures produced either from GaAs or GaAsP. After 8 years in 1970 was obtained the first room temperature continuous wave operation by Zh. Alferov et al. using the double heterostructure on the AlGaAs/GaAs material system [2]. The following progress due to the inclusion of quantum wells make possible to decrease the threshold current density from 1000 A/cm\textsuperscript{2} range to the 10 A/cm\textsuperscript{2} range [2]. All those “mainstream” semiconductor lasers with rather high photon energy of electronvolt range use the band-to-band electron-hole recombination principle and are relatively simple to design.

Idea of alternative semiconductor lasers, the quantum cascade lasers (QCL), based on sophisticated alteration of the band gap parameters in nanoscale heterostructures was proposed by Kazarinov and Suris in 1971 [3]. The first QCL of mid-infrared range was realized in 1994 by the Faist et al. (see references within [4]). The first QCL of greater wavelengths of terahertz range was fabricated in 2002 by Koler et al [4]. Important is to notice the 1998 paper of Sirtori, Faist et al. [5] where the “classic” heterostructure compound AlGaAs/GaAs with excellent lattice match properties was first used also for QCL. In general, the longer wavelengths of terahertz range (\(\lambda > 40\ \mu \text{m}\)) are more difficult to realize [4] and need more sophisticated design tools.

The present note discusses the development and optimization of GaAs/AlGaAs QCL simulation software that has been described in three editions of the book [6].

2 Development guidelines

The software of Leeds University [6] has been developed for scientific research purposes with high level of autonomy of invoked submodules. The computing platform is the multi-processor cluster. The used universality causes, however, a remarkable recalculation of physical quantities in multi-level hierarchical calling structures. For instance, in study of the spontaneous THz radiation source [7], nearly 2 orders of magnitude acceleration was achieved due to the application specific modification of software. Thus in present study the goal was to test the feasibility of development of the desktop QCL simulator with calculation time of 10 minute range.

The QCL simulation algorithm contains in general lines of the following steps [6, 7]:
1. Fix the structure (barriers and wells) for one cascade period;
2. Fix the bias electric field \(F\);
2a. Solve Schrödinger equation for one cascade period;
2b. Solve Schrödinger equation for three cascade periods;
2c. Form the periodic solution for 2 adjacent periods.
3. Fix the lattice temperature \(T_l\);
4. Fix electron populations \(n_i\) for all states (sum defined by doping, initially uniform, 1st and 2nd period equal);
5. Fix or estimate the common electron temperature \(T_e\);
6. Calculate Fermi energy levels by the known electron populations (concentrations) for every state;
7. Calculate raw scattering rates (wavevector-dependent) for all scattering mechanism (presently six [6]);
8. Find averaged scattering rates for all transitions between all states of 2 periods;
9. Form the rate equations system and solve it to find corrected populations \(n_i\);
10. If populations are not established, return to point 6;
11. If wanted, try to modify electron temperature \(T_e\);
12. If \(T_e\) is not established, return to 5;
13. Calculate current density, gain and other output.

978-1-4673-2774-9/12/$31.00 ©2012 IEEE
The proposed acceleration methodology for development of the optimized simulator include:
- avoid repeated recalculation of physical quantities;
- analyze the optimal order of integration loops;
- use automatic storing of intermediate quantities in tables instead of repeated recalculation;
- use analytic approximations to rise the integration accuracies at smaller number of argument steps;
- use more data connections via the RAM instead of total file-based connections.

In further development the additional feedback loop due to the photon field intensity and stimulated emission in active layer of QCL should be considered.

3 Example results

Fig. 1 demonstrates the minimal QCL structure containing three quantum wells (three energy levels or quantum states) per period [6]. This structure may be used for testing of simulator before discussing more complex structures, e.g. with 8 states per period [5].

![Figure 1 Coordinate-dependencies of the electron potential energy and the wavefunction squares (in arbitrary units, level energies [meV] added) in two adjacent periods of quantum cascade laser structure under applied voltage. The lasing transition is 3→2 (also 6→5 due to similarity). The “cleaning” transition 2→1 must be faster (i.e. with higher scattering rate) than 3→2 in order to achieve population inversion \( n_3 > n_2 \). Transition 1→6 must be also fast to assure high electron concentration on the upper lasing level of the next period.](image1)

Fig. 2 analyses the conditions of laser radiation generation in this structure: ratio of transition times 3→2 versus 2→1 and population ratio between levels 3 and 2. High transition times ratio \( \tau_{32}/\tau_{21} \) is a precondition for laser operation but actual simulation shows that the necessary inversion of population \( n_3/n_2 > 1 \) appears only between electric field values 12.5 and 17 kV/cm.

Fig. 3 shows the corresponding current density through the device and the possible light output (in arbitrary units as the real output depends on number of periods, optical resonator parameters etc.).

The introduced software optimization measures gave approximately 10-fold calculation speed enhancement to the range of 1-2 minutes per one field value on Intel Core 2 Duo desktop PC (depending on spatial grid size).

![Figure 2 Ratio of lasing (3→2) and cleaning (2→1) transition times and ratio of electron populations (concentrations) on upper (3) and lower (2) laser levels versus applied field.](image2)

![Figure 3 Current density and expected laser radiation output (in arbitrary units) versus applied field.](image3)

References

Curriculum vitae

1. Personal data

| Name: | Reeno Reeder |
| Date and place of birth: | 07/12/1981, Tallinn |
| E-mail address: | reeno.reeder@ttu.ee |

2. Education

<table>
<thead>
<tr>
<th>Educational institution</th>
<th>Graduation year</th>
<th>Education (field of study/degree)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tallinn University of Technology</td>
<td>2005</td>
<td>Electronics and Biomedical Engineering/BSc</td>
</tr>
<tr>
<td>Tallinn University of Technology</td>
<td>2007</td>
<td>Electronics and Biomedical Engineering/MSc</td>
</tr>
<tr>
<td>Tallinn University</td>
<td>2013</td>
<td>Teacher's Professional Studies</td>
</tr>
</tbody>
</table>

3. Language competence/skills

<table>
<thead>
<tr>
<th>Language</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estonian</td>
<td>Fluent</td>
</tr>
<tr>
<td>English</td>
<td>Fluent</td>
</tr>
<tr>
<td>Russian</td>
<td>Basic skills</td>
</tr>
<tr>
<td>Swedish</td>
<td>Basic skills</td>
</tr>
</tbody>
</table>

4. Special courses

<table>
<thead>
<tr>
<th>Period</th>
<th>Educational or other organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/09/2006 - 22/12/2006</td>
<td>Visitor student in Leeds University</td>
</tr>
<tr>
<td>30/06/2012 - 21/07/2012</td>
<td>Participation in CERN International High School Teachers Programme</td>
</tr>
</tbody>
</table>

5. Professional employment

<table>
<thead>
<tr>
<th>Period</th>
<th>Organisation</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/05/2002 - 31/07/2011</td>
<td>Cybernetica AS</td>
<td>Programmer, System Analyst</td>
</tr>
<tr>
<td>22/08/2011 - 31/08/2013</td>
<td>Keila Hariduse sihtasutus</td>
<td>Physics Teacher</td>
</tr>
<tr>
<td>01/08/2013 - ...</td>
<td>Tallinn University of Technology</td>
<td>Teaching Assistant</td>
</tr>
</tbody>
</table>
6. Research activity

Honours:
Students' research diploma by Archimedes in 2005.

Theses:
Master thesis: Modelling of terahertz broadband radiation sources based on GaAs/AlGaAs quantum well heterostructures (supervised by Andres Udal, 2007)

Publications:


Curriculum vitae

1. Isikuandmed

Nimi: Reeno Reeder
Sünniaeg ja koht: 07/12/1981, Tallinn
Kodakondsus: Eesti
E-posti aadress: reeno.reeder@ttu.ee

2. Hariduskaik

<table>
<thead>
<tr>
<th>Õppeasutus</th>
<th>Lõpetamise aeg</th>
<th>Haridus (eriala/kraad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tallinna Tehnikaülikool</td>
<td>2005</td>
<td>elektroonika ja biomeditsüinitehnika/BSc</td>
</tr>
<tr>
<td>Tallinna Tehnikaülikool</td>
<td>2007</td>
<td>elektroonika ja biomeditsüinitehnika/MSc</td>
</tr>
<tr>
<td>Tallinna Ülikool</td>
<td>2013</td>
<td>õpetaja kutseõpingud</td>
</tr>
</tbody>
</table>

3. Keelteoskus

<table>
<thead>
<tr>
<th>Keel</th>
<th>Tase</th>
</tr>
</thead>
<tbody>
<tr>
<td>eesti keel</td>
<td>kõrgtase</td>
</tr>
<tr>
<td>inglise keel</td>
<td>kõrgtase</td>
</tr>
<tr>
<td>vene keel</td>
<td>algtase</td>
</tr>
<tr>
<td>rootsi keel</td>
<td>algtase</td>
</tr>
</tbody>
</table>

4. Täiendusõpe

<table>
<thead>
<tr>
<th>Õppimise aeg</th>
<th>Täiendusõppe korraldaja nimetus</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/09/2006 - 22/12/2006</td>
<td>Külalistudeng Leedsi Ülikoolis</td>
</tr>
<tr>
<td>30/06/2012 - 21/07/2012</td>
<td>Osalemine CERN-i rahvusvahelises gümnaasiumiõpetajate programmis</td>
</tr>
</tbody>
</table>

5. Teenistuskaik

<table>
<thead>
<tr>
<th>Töötamise aeg</th>
<th>Tööandja nimetus</th>
<th>Ametikoht</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/05/2002 - 31/07/2011</td>
<td>Cybernetica AS</td>
<td>programmeerija, süsteemianalüütik</td>
</tr>
<tr>
<td>22/08/2011 - 31/08/2013</td>
<td>Keila Hariduse sihtasutus</td>
<td>füüsikaõpetaja</td>
</tr>
<tr>
<td>01/08/2013 - ...</td>
<td>Tallinna Tehnikaülikool</td>
<td>assistent</td>
</tr>
</tbody>
</table>
6. Teadustegevus

Tunnustused:
Diplom üliõpilaste teadustööde riiklikult konkursilt, Archimedes, 2005.

Lõputööd:
Bakalaureusetöö: Sissejuhatus kvantmehaanika numbrilistesse meetoditesse (juhendaja Andres Udal, 2005)
Magistritöö: Kvantukudega GaAs/AlGaAs heterostruktuuridel põhinevate laiaribaliste terahertskiirgurite modelleerimine (juhendaja Andres Udal, 2007)

Publikatsioonid:


DISSERTATIONS DEFENDED AT
TALLINN UNIVERSITY OF TECHNOLOGY ON
INFORMATICS AND SYSTEM ENGINEERING

4. **Kaido Kikkas.** Using the Internet in Rehabilitation of People with Mobility Impairments – Case Studies and Views from Estonia. 1999.


30. Виктор Войтович. Разработка технологий выращивания из жидкой фазы эпитаксиальных структур арсенида галлия с высоковольтным р-п переходом и изготовления диодов на их основе. 2006.


40. **Andrei Pokatilov.** Development of National Standard for Voltage Unit Based on Solid-State References. 2008.


42. **Maksim Jenihhin.** Simulation-Based Hardware Verification with High-Level Decision Diagrams. 2008.


47. **Andrei Krivošei.** Model Based Method for Adaptive Decomposition of the Thoracic Bio-Impedance Variations into Cardiac and Respiratory Components. 2009.

48. **Vineeth Govind.** DfT-Based External Test and Diagnosis of Mesh-like Networks on Chips. 2009.


51. **Erika Matsak.** Discovering Logical Constructs from Estonian Children Language. 2009.


57. **Pavel Grigorenko.** Higher-Order Attribute Semantics of Flat Languages. 2010.


82. **Igor Aleksejev.** FPGA-based Embedded Virtual Instrumentation. 2013.


89. **Jaas Ježov.** Pressure sensitive lateral line for underwater robot. 2013.

90. **Vadim Kaparin.** Transformation of Nonlinear State Equations into Observer Form. 2013.