GaAs/AlGaAs Far-Infrared Quantum Cascade Laser

by

Hans Callebaut

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science
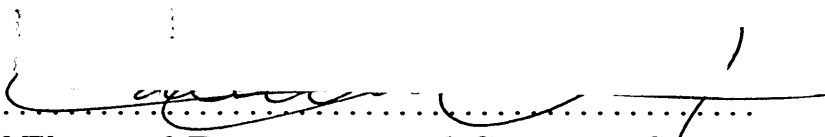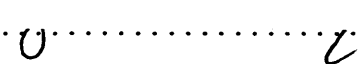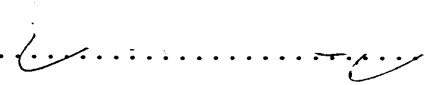
at the

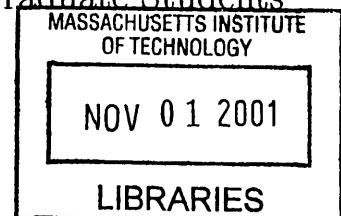MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2001
[September 2001]

Author ....
Department of Electrical Engineering and Computer Science
August 6, 2001

Certified by...
Qing Hu
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

BARKER

GaAs/AlGaAs Far-Infrared Quantum Cascade Laser

by

Hans Callebaut

Submitted to the Department of Electrical Engineering and Computer Science
on August 6, 2001, in partial fulfillment of the
requirements for the degree of
Master of Science in Electrical Engineering and Computer Science

Abstract

In this thesis I investigated the feasibility of an optically pumped intersubband far-infrared (40-100 μm) laser, using GaAs/Al$_x$Ga$_{1-x}$As heterostructures. The proposed design aims to use LO-phonon-mediated depopulation of the lower THz laser level to aid the intersubband laser population inversion. Interband recombination occurs by means of stimulated emission, thus combining an interband ($\sim$ 1550 meV) and intersubband ($\sim$ 16-18 meV) laser.

As the subband properties of both the valence band and the conduction band are important for this work, a numerical program code was developed for the valence band to supplement the available tools for the conduction band. The steady state rate equations for the proposed quantum well structure were solved self-consistently for several different carrier temperatures. The calculations indicate that a pump beam of moderate power (0.5-1 W) concentrated on a device of typical dimensions (10^4 cm^2) can generate an intersubband gain of 20 cm^{-1} at 50 K for a THz emission linewidth of 2 meV. This gain level can suffice to obtain THz lasing action, provided that the cavity losses can be kept in check. The performance of the THz laser is predicted to be very dependent on electron temperature, mainly due to the opening of a parasitic LO-phonon channel between the THz laser levels. Interband lasing seems to be easier to obtain, as the calculated threshold pump intensity is lower than for the intersubband case.

Thesis Supervisor: Qing Hu
Title: Associate Professor of Electrical Engineering and Computer Science

Acknowledgments

Contents

List of Figures

List of Tables

.

Chapter 1

Introduction

1.1 Interband Pumped Intersubband THz Lasers

1.1.1 Introduction

The far-infrared frequency range is roughly defined as 30-300 μm or 4-40 meV. Often this range is also referred to by the term terahertz radiation, since 4-40 meV corresponds to 1-10 THz. Far-infrared (FIR) or terahertz (THz) electromagnetic radiation is important in many applications such as radio astronomy, environmental monitoring, plasmon diagnostics, laboratory spectroscopy, telecommunications etc. and in the characterization of nanoscale condensed matter materials. In recent years, the generation, propagation and detection of FIR or THz electromagnetic radiation using two-dimensional semiconductor systems or other semiconductor nanostructures has become one of the most rapidly expanding fields in the photonics, optoelectronics and condensed matter physics communities.

Diode lasers are ideal sources because they are cheap, compact and very efficient. However, the semiconductor band gap places a limitation on emission frequency. The longest-wavelength diode lasers ($\sim 30\mu$m) are based on narrow gap lead-salt semiconductors [1]. While these lead-salt lasers have been quite successful for high resolution

spectroscopy, they are still limited to cryogenic operation and provide relatively low power. On the other end of the spectrum, semiconductor transistors can be used to make 100 GHz oscillators [2]. Molecular gas lasers are currently the only practical laser sources for the far infrared, but they have limited lasing frequencies. They are also somewhat unwieldy as they require high voltage supplies and are usually rather bulky.

Intersubband lasers have several advantages over conventional semiconductor lasers. Most useful is the fact that the emission frequency is chosen by the design of the widths of the quantum wells, and can hence be tailored to the application. This is especially useful for infrared applications where small bandgap materials become difficult to find and work with. Also, since the envelope functions extend over a well (tens of Angstroms), the dipole moment for the intersubband transition is typically several orders of magnitude larger than that of an atomic transition. These features promise more efficient lasers.

1.1.2   Intersubband Lasers

In 1970, Esaki and Tsu [3] proposed using heterostructures for applications in opto-electronics. The use of intersubband transitions to create a laser was first suggested by Kazarinov and Suris [4] in 1971. Since then, electrically pumped quantum cascade lasers have been developed for wavelengths up to 24 μm [5]. Quantum wells are made by growing layers of different band gap semiconductors on top of each other, creating a stack like structure. Since the bandgap of GaAs is smaller than that of $Al_xGa_{1-x}As$, the ensuing band gap profile gives rise to potential wells. The potential well height is determined by the Al alloy concentration of the barrier material.

These quantum wells perturb the crystal periodicity in the growth direction. New electron energy states are located in these quantum wells, confined in the growth direction but still free in the plane of the well. As shown in figure (1-1), the conduction band is quantized into subbands.
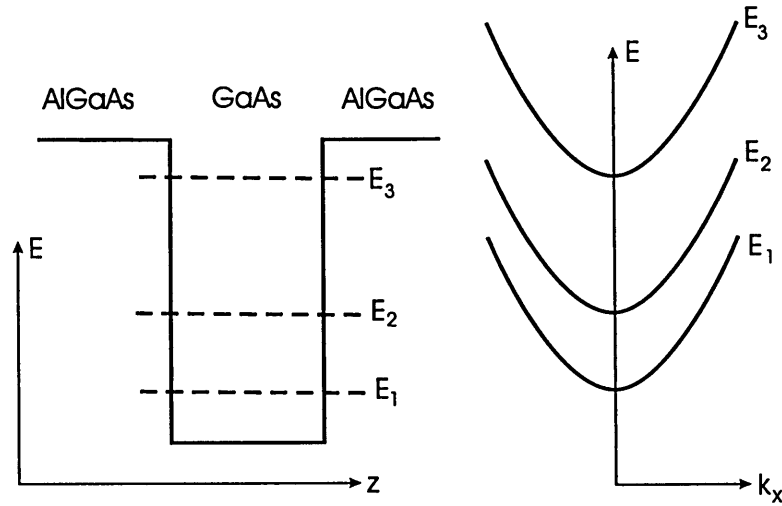
Figure 1-1: *Subbands in a quantum well. The potential well caused by the AlGaAs / GaAs quantum well gives rise to bound states localized in the well. In k-space, the subbands are parabolic as the electrons are not confined in the plane of the well.*

The quantum well is similar to an impurity atom in that localized states are created. In k-space, the subbands are parabolic as the electrons are not confined in the plane of the well. The exact energies of the subband minima are dependent on the well width and the depth of the potential well. The energies can be approximated by the formula for infinitely deep wells:

$$E_n = \frac{\hbar^2}{2m^*}(\frac{n\pi}{L})^2, \tag{1.1}$$

where m^* is the electron effective mass in GaAs, L is the well width and $\hbar$ is the reduced Planck's constant. The energy levels for a well with a finite barrier are lower than in equation (1.1). By choosing the well width and the barrier heights we can tailor the quantum levels so that transitions between E_m and E_n will emit photons in the far infrared. Very interesting from an engineering point of view is the ability to tune the energy levels and the dipole moments by applying a voltage bias. The Stark shift induced by the electric field shifts the energy levels and alters the potential profile. This is a very powerful tool when designing quantum wells.
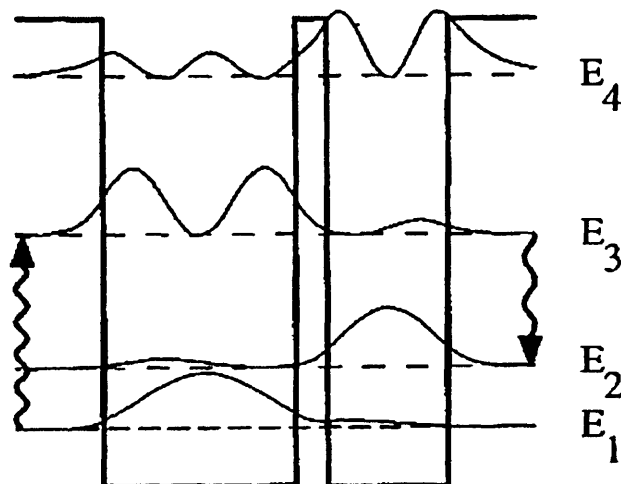
Figure 1-2: *Gauthier-Lafaye's three level quantum well design. Indicated are the pump and mid-IR intersubband laser transitions.* [8]

Kazarinov and Suris [4] were the first to propose the use of intersubband transitions to design a laser. Electrically pumped versions of such a laser, called a Quantum Cascade Laser, have been demonstrated in recent years. The QCL works by applying a voltage bias to create a potential staircase in which each step consists of a module of quantum wells. In most designs, each module has three major energy levels. Electrons are injected into the upper level and relax down to the middle level by emission of a photon. Subsequently, this lower laser level is quickly vacated by fast polar longitudinal optical (LO) phonon scattering to the ground state. From this ground state, electrons are then injected into the upper level of the next module. This design and variations thereof have been very successful in the mid-infrared frequency range 4-24 μm [5, 6, 7].

An alternative to electrical pumping is optical pumping, either intersubband or interband. Room temperature luminescence around 10 μm was observed by Sauvage *et al* [9] in 1996 and at 7.7 μm in 1997 [10]. Lyobumirsky *et al.* reported spontaneous
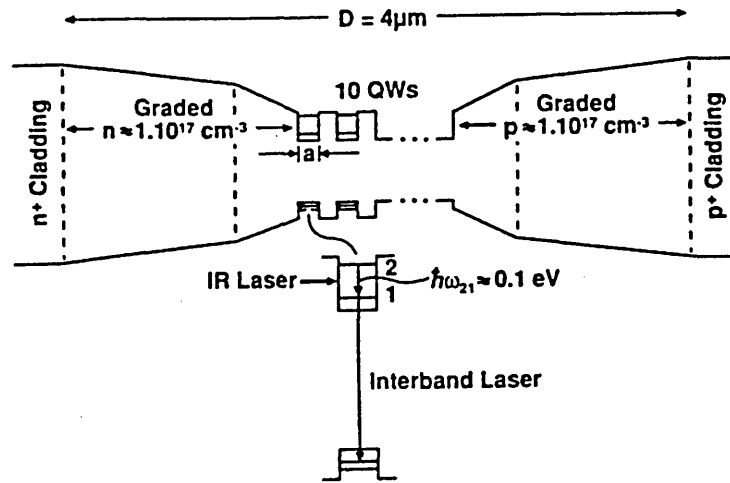
Figure 1-3: *Kastalsky's MQW design consists of multiple single wells. Also shown is a scheme of the optical transitions under IR laser operation [14].*

far-infrared emission around 47 μm [11]. In 1998, a mid infrared quantum well laser operating at 15.5 μm was realized by Gauthier-Lafaye *et al.* [8]. A schematic diagram of their laser design is shown in figure (1-2). This laser uses a simple three-level scheme where electrons are optically pumped from the ground level into the excited state E_3. Population inversion is ensured by designing the intersubband spacing $E_2 - E_1 \approx \hbar\omega_{LO}$ for LO phonon resonance to depopulate E_2. Similar structures were proposed by Julien [12], Green [13] and Sauvage [10].

Optical pumping can have the advantage of a high selectivity in populating energy levels in a simple structure. It avoids the free carrier losses associated with contact regions. Also, a voltage bias can be seen as an extra degree of freedom and is not instrumental in carrier transport between modules. On the other hand, electrical pumping is a more convenient and efficient.

Recently, several proposals and measurements of FIR interband pumped laser devices have been published, but no lasing has been observed yet. Sauvage *et al.* [10, 9] reported both far- and mid-infrared intersubband luminescence in an interband

pumped device. Kastalsky [14] proposed an infrared intersubband laser based on an interband laser. In his design a pump laser excites electrons into the first excited state in an isolated quantum well. The ground state is depopulated by a stimulated interband process (figure (1-3)). Vurgaftman and Meyer proposed the use of optically pumped type-II heterostructures to obtain THz lasing [15].

For intersubband pumping, the CO_2 laser is the most commonly used laser. Its main drawback is that its wavelength is fixed, which imposes limits on the QW design. For interband pumped devices, the Ti:Sapphire laser offers easy tunability from 690-1080 nm wavelength and a range of pulse lengths continuously variable from 80 ps to less than 50 fs. Also, cheap single mode laser diodes around 780 nm are available with power levels up to 100 mW. SLI Corporation (15 Link Drive, Binghamton, NY 13904) offers a single mode diode laser at 785 nm (e.g. SLI-CW-XXX-C1-XXX-0.1S-R) with up to 100 mW output power. Multimode laser diodes and diode laser arrays (linewidth around 3 nm) are more commonly available with powers ranging to upwards of 200 W. Coherent offers a single-stripe multimode laser diode with 3 W output power (S-79-3000C-200x) at 780 nm. The same company offers a 20 W laser diode array (B1-770-20C-19-30-A) at 770 nm. Unfortunately it is very difficult to find laser diodes with a wavelength between 740 and 780 nm (1670 and 1600 meV).

The proposed pumping geometry is based on the setup used by Le *et al.* [16] We can use a linear pump diode array with a collimating cylindrical lens. The collimated, parallel beam can be focused into a stripe by an aspheric condenser. The stripe width can be controlled with a pair of wedges that form a precise slit aperture and also act as a beam scrambler to reduce hot spots. The wedges are positioned slightly off the condenser focus so that the pump density profile along the stripe is nearly uniform. The sample itself can be mounted on a cold finger inside a dewar. Intersubband radiation emitted by the sample is coupled out of the dewar through a window. In order to ensure an efficient use of the pumping power, we will assume surface pumping. This means that the incident pump beam propagates along the QW growth axis, and
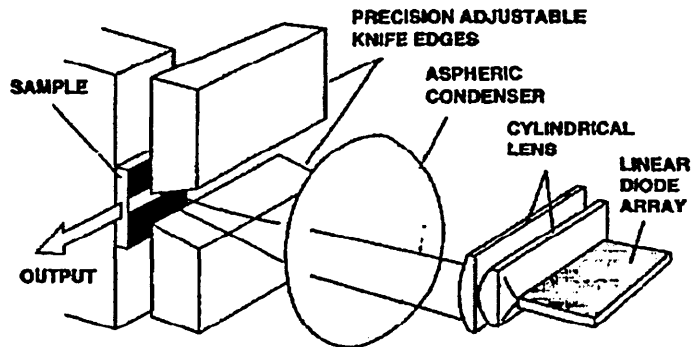
Figure 1-4: *Schematic drawing of the optical pumping arrangement [16]*

the electric field is polarized in the plane of the well (TE polarization). The active region is surrounded by cladding layers and a metal-metal waveguide. The top metal layer should have windows to allow the pump beam to pass through, yet still reflect the THz emission for an effective mode confinement. This can be achieved by having the rectangular windows aligned with their length along the ridge. Edge pumping is more difficult because of the very small penetration depth of the pump field in the sample (on the order of 1 μm). On top of that, it is difficult to focus the pump beam exactly on the sample's edge.

The proposed interband pumped intersubband laser consists of a sequence of independent modules of three coupled quantum wells of increasing width. A laser emitting radiation of the appropriate frequency excites carriers in the narrowest well (containing the highest energy ground state) from the valence band into the corresponding conduction subband which will serve as the FIR laser's upper level. The middle well, of intermediate width, hosts the lower laser level. Finally, this lower laser level is depopulated by making use of resonant LO phonon scattering into the third and widest well. In order to depopulate this well and avoid state blocking for the LO-phonon transition from the lower laser level, interband lasing conditions must be met.

1.2   Problem Statement

Two major difficulties pose challenges to the design of interband pumped intersub-band THz lasers: pump selectivity and the lasing conditions for the "depopulation laser" in the widest well. Ideally, the pump laser would only excite carriers in the narrowest well, containing the upper laser level. However, as this well also has the largest "effective bandgap" of all three wells in a module, it is inevitable that the pump laser energy will be greater than the effective bandgaps in the other two wells, giving rise to an undesirable generation of photo-excited carriers in these wells. Those carriers reduce the population inversion between the intersubband laser levels by swamping the lower FIR laser level with electrons, making population inversion more difficult to obtain. Also, by increasing the carrier density in the widest well, state blocking of the depopulation of the lower level may become an issue. Furthermore, this "parasitic" photogeneration depletes the pump beam, decreasing the attainable efficiency of the device. Secondly, in order to determine the depopulation laser threshold, it is necessary to investigate the hole transport in the valence band, as well as the interband dipole moment.

Chapter 2

Optical Transitions

2.1 Introduction

In order to evaluate electronic and optical processes in a semiconductor, an adequate model of its electronic band structure is needed, defining electron (and hole) energy levels and effective masses. This will enable us to calculate the corresponding wave functions, and hence absorption and gain due to electronic transitions in the presence of an incident optical field. Most important for optical devices are optical transitions between conduction and valence band edge states, which are lowest in energy and most densely populated. Therefore, we will focus on the conduction and valence band structures near the band edge, where the $\mathbf{k} \cdot \mathbf{p}$ method is very useful.

2.2 Bloch Theorem and $\mathbf{k} \cdot \mathbf{p}$ Method

A crystal lattice is characterized by its long range order and symmetry. This symmetry will be reflected in the crystal's electrostatic potential, which in turn influences the movement of charge carriers. How these (translation) symmetry properties can be expressed in terms of properties of the carrier wave functions, is described by the Bloch Theorem.

For an electron in a periodic potential $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$, with $\mathbf{R}$ a lattice vector, the electron wave function satisfies the Schrödinger equation :

$$H\psi(\mathbf{r}) = \left[\frac{-\hbar^2}{2m_0}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E(\mathbf{k})\psi(\mathbf{r}). \qquad (2.1)$$

The Hamiltonian is invariant under the translation $\mathbf{r} \to \mathbf{r} + \mathbf{R}$. Therefore, if $\psi(\mathbf{r})$ is a solution, $\psi(\mathbf{r} + \mathbf{R})$ will also satisfy (2.1), the only possible difference being a constant phase factor (which has no physical meaning). So, the general solution to (2.1) can be written as:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}), \qquad (2.2)$$

with

$$u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = u_{n\mathbf{k}}(\mathbf{r}), \qquad (2.3)$$

n referring to the band and $\mathbf{k}$ to the electron wave vector.

The $\mathbf{k} \cdot \mathbf{p}$ method is a useful technique for analyzing the band structure near a particular point $\mathbf{k_0}$, especially when it is near a band extremum. Here, we will consider the case where this extremum occurs at the zone center, $\mathbf{k_0} = 0$. This is a very useful case for direct III-V semiconductors, such as GaAs.

Substituting (2.2) in (2.1) we can write the Schrödinger equation in terms of $u_{n\mathbf{k}}(\mathbf{r})$:

$$\left[\frac{p^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p} + V(\mathbf{r})\right]u_{n\mathbf{k}}(\mathbf{r}) = \left[E_n(\mathbf{k}) - \frac{\hbar^2 k^2}{2m_0}\right]u_{n\mathbf{k}}(\mathbf{r}), \qquad (2.4)$$

or

$$[H_0 + H']u_{n\mathbf{k}}(\mathbf{r}) = \mathcal{E}_n(\mathbf{k})u_{n\mathbf{k}}(\mathbf{r}), \qquad (2.5)$$

where

$$H_0 = \frac{p^2}{2m_0} + V(\mathbf{r}), \qquad (2.6)$$

$$H' = \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p}, \qquad (2.7)$$

$$\mathcal{E}_n(\mathbf{k}) = E_n(\mathbf{k}) - \frac{\hbar^2 k^2}{2m_0}. \tag{2.8}$$

If we're just interested in a single band, as is the case near the band edge in a conduction band, we can easily use perturbation theory to calculate the energy dispersion relation. We carry out an expansion around $\mathbf{k} = 0$, and time independent perturbation theory gives the energy to second order [17]:

$$E_n(\mathbf{k}) = E_n(0) + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} \mathbf{k}\cdot\mathbf{p}_{nn} + \frac{\hbar^2}{m_0^2} \sum_{n'\neq n} \frac{|\mathbf{k}\cdot\mathbf{p}_{nn'}|^2}{E_n(0) - E_{n'}(0)}, \tag{2.9}$$

where the momentum matrix elements are defined as

$$\mathbf{p}_{nn'} = \int_{\substack{unit\\cell}} u_{n0}^\dagger(\mathbf{r})\mathbf{p}u_{n'0}(\mathbf{r})d^3\mathbf{r}. \tag{2.10}$$

If $\mathbf{k}_0$ is at an extremum of $E_n(\mathbf{k})$, $\mathbf{p}_{nn}$ usually vanishes because of symmetry considerations, and $E_n(\mathbf{k})$ around the extremum has a second order dependence on the components of $\mathbf{k}$. We can write this as :

$$E_n(\mathbf{k}) - E_n(0) = \sum_{\alpha,\beta} D^{\alpha\beta} k_\alpha k_\beta = \frac{\hbar^2}{2} \sum_{\alpha,\beta} \left(\frac{1}{m^*}\right)_{\alpha\beta} k_\alpha k_\beta, \tag{2.11}$$

and

$$D^{\alpha\beta} = \frac{\hbar^2}{2m_0}\delta_{\alpha\beta} + \frac{\hbar^2}{2m_0^2} \sum_{n\neq n'} \frac{p_{nn'}^\alpha p_{n'n}^\beta + p_{nn'}^\beta p_{n'n}^\alpha}{E_n(0) - E_{n'}(0)} = \frac{\hbar^2}{2} \left(\frac{1}{m^*}\right)_{\alpha\beta}, \tag{2.12}$$

where α, $\beta = x$, y and z. The matrix $D^{\alpha\beta}$ is the inverse effective mass in matrix form multiplied by $\hbar^2$

2.3 Luttinger-Kohn's model

2.3.1 Löwdin's Perturbation Method

These results from the single-band $\mathbf{k}\cdot\mathbf{p}$ theory can be generalized for the case of multiple, degenerate bands by using the Luttinger-Kohn model [18, 19]. In this model the influence of all other bands are taken into account by using Löwdin's perturbation method [17]. All bands are subdivided into two classes:
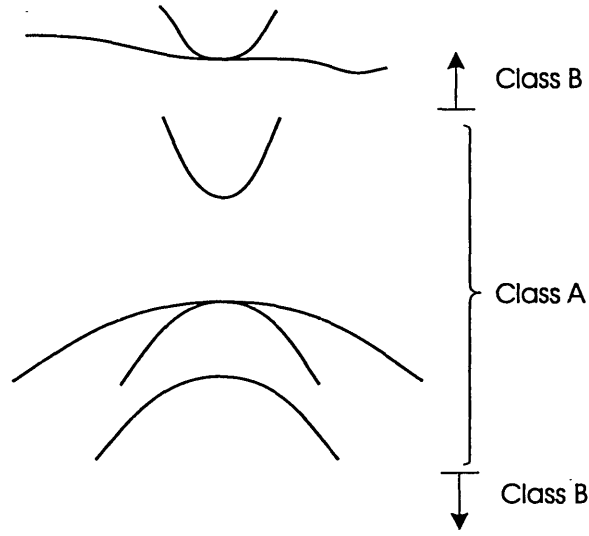
Figure 2-1: *Conduction and valence bands are divided into two classes for the application of Löwdin's perturbation method.*

- class A : the six valence bands(heavy hole, light hole, split off band and their spin counterparts) and the two conduction bands

- class B : all other bands

We will concentrate on the bands in class A, while taking into account class B bands perturbatively.

We can write the perturbed solution ϕ as a linear combination of the unperturbed eigenstates $\phi_i^{(0)}$:

$$\phi = \sum_n^{A,B} a_n \phi_n^{(0)}. \tag{2.13}$$

Assuming the unperturbed eigenstates are orthonormalized, we can write the eigenequation as:

$$(E - H_{mm})a_m = \sum_{n \neq m}^{A} H_{mn}a_n + \sum_{\alpha \neq m}^{B} H_{m\alpha}a_\alpha, \tag{2.14}$$

with

$$H_{mn} = \int \phi_m^{(0)\dagger} H \phi_n^{(0)} d^3\mathbf{r} = E_n^{(0)}\delta_{mn} + H'_{mn}, \tag{2.15}$$

or

$$a_m = \sum_{n \neq m}^{A} \frac{H_{mn}}{E - H_{mm}} a_n + \sum_{\alpha \neq m}^{B} \frac{H_{m\alpha}}{E - H_{mm}} a_\alpha, \qquad (2.16)$$

where the first sum on the right-hand side is over the states in class A only, while the second sum is over the states in class B. Since we are interested in the coefficients a_m for m in class A, we may eliminate those in class B by an iteration procedure and obtain:

$$a_m = \sum_{m \neq n}^{A} \frac{U_{mn}^{A} - H_{mn}\delta_{mn}}{E - H_{mm}} a_n, \qquad (2.17)$$

and

$$U_{mn}^{A} = H_{mn} + \sum_{\alpha \neq m}^{B} \frac{H_{m\alpha}H_{\alpha n}}{E - H_{\alpha\alpha}} + \sum_{\substack{\alpha, \beta \neq m, n \\ \alpha \neq \beta}} \frac{H_{m\alpha}H_{\alpha\beta}H_{\beta n}}{(E - H_{\alpha\alpha})(E - H_{\beta\beta})} + \cdots \qquad (2.18)$$

Or, equivalently, we solve the eigenvalue problems for a_n, $(n \in A)$

$$\sum_{n}^{A} \left(U_{mn}^{A} - E\delta_{mn} \right) a_n = 0 \qquad m \in A, \qquad (2.19)$$

and

$$a_\gamma = \sum_{n}^{A} \frac{U_{\gamma n}^{A} - H_{\gamma n}\delta_{\gamma n}}{E - H_{\gamma\gamma}} a_n \qquad \gamma \in B. \qquad (2.20)$$

When the coefficients a_n belonging to class A are determined from the eigenequation (2.19), the coefficients a_γ can be found from (2.20). A necessary condition for the expansion of (2.18) is:

$$|H_{m\alpha}| \ll |E - H_{\alpha\alpha}| \qquad m \in A, \ \alpha \in B. \qquad (2.21)$$

2.3.2   Schrödinger Equation and Basic Functions

Including the spin-orbit interaction, the Hamiltonian is written as [17]:

$$H = H_0 + \frac{\hbar}{4m_0^2 c^2} \overline{\sigma} \cdot \nabla V \times \mathbf{p}, \tag{2.22}$$

where $\overline{\sigma}$ is the Pauli spin matrix vector. Substituting 2.2 in the Schrödinger equation, we finally obtain

$$
\begin{aligned}
H u_{n\mathbf{k}}(\mathbf{r}) &= \left(H_0 + \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{\Pi} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{4m_0^2 c^2} \nabla V \times \mathbf{p} \cdot \overline{\sigma} \right) u_{n\mathbf{k}}(\mathbf{r}) \\
&= E_n(\mathbf{k}) u_{n\mathbf{k}}(\mathbf{r}),
\end{aligned}
\tag{2.23}
$$

where

$$\mathbf{\Pi} = \mathbf{p} + \frac{\hbar}{4m_0 c^2} \overline{\sigma} \times \nabla V, \tag{2.24}$$

and the perturbation Hamiltonian

$$H' = \frac{\hbar}{m_0} \mathbf{k} \cdot \mathbf{\Pi}. \tag{2.25}$$

The unperturbed Hamiltonian will then refer to the band-edge spin-orbit system (for $k = 0$).

Note that the second term in $\mathbf{k} \cdot \mathbf{\Pi}$ is much smaller than the last term in (2.23), because $\hbar k \ll p = |\langle u_k | \mathbf{p} | u_k \rangle| \approx \hbar/a$. The electron velocity in the atomic orbit is much larger than the velocity of the wave packet in the vicinity of $\mathbf{k}_0 \approx 0$ (band edge).

At the band edge, conduction band Bloch states exhibit s-like symmetry, and valence band states are p-like (3-fold degenerate without spin). We can represent these states by $|S\rangle$, and $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ respectively. We can picture these Bloch functions as a periodic repetition of atomic orbitals, repeated at intervals corresponding to the lattice spacing.

We expand the function

$$u_{nk}(\mathbf{r}) = \sum_{j'}^{A} a_{j'}(\mathbf{k})u_{j'0}(\mathbf{r}) + \sum_{\gamma}^{B} a_{\gamma}(\mathbf{k})u_{\gamma0}(\mathbf{r}), \qquad (2.26)$$

where j' is in class A and γ is in class B. We choose the basic functions $u_{j'0}$ to be

$$u_{10}(\mathbf{r}) = u_{el}(\mathbf{r}) = \left|S\frac{1}{2}, \frac{1}{2}\right\rangle = |S\uparrow\rangle,$$

$$u_{20}(\mathbf{r}) = u_{SO}(\mathbf{r}) = \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{\sqrt{3}}|(X+iY)\downarrow\rangle + \frac{1}{\sqrt{3}}|Z\uparrow\rangle,$$

$$u_{30}(\mathbf{r}) = u_{lh}(\mathbf{r}) = \left|\frac{3}{2}, \frac{1}{2}\right\rangle = \frac{-1}{\sqrt{6}}|(X+iY)\downarrow\rangle + \sqrt{\frac{3}{2}}|Z\uparrow\rangle,$$

$$u_{40}(\mathbf{r}) = u_{hh}(\mathbf{r}) = \left|\frac{3}{2}, \frac{3}{2}\right\rangle = \frac{-1}{\sqrt{2}}|(X+iY)\uparrow\rangle,$$

$$u_{50}(\mathbf{r}) = \overline{u}_{el}(\mathbf{r}) = \left|S\frac{1}{2}, \frac{-1}{2}\right\rangle = -|S\downarrow\rangle,$$

$$u_{60}(\mathbf{r}) = \overline{u}_{SO}(\mathbf{r}) = \left|\frac{1}{2}, \frac{-1}{2}\right\rangle = \frac{1}{\sqrt{3}}|(X-iY)\uparrow\rangle - \frac{1}{\sqrt{3}}|Z\downarrow\rangle,$$

$$u_{70}(\mathbf{r}) = \overline{u}_{lh}(\mathbf{r}) = \left|\frac{3}{2}, \frac{-1}{2}\right\rangle = \frac{1}{\sqrt{6}}|(X-iY)\uparrow\rangle + \sqrt{\frac{3}{2}}|Z\downarrow\rangle,$$

$$u_{80}(\mathbf{r}) = \overline{u}_{hh}(\mathbf{r}) = \left|\frac{3}{2}, \frac{-3}{2}\right\rangle = \frac{1}{\sqrt{2}}|(X-iY)\downarrow\rangle. \qquad (2.27)$$

Using Löwdin's method, we need only solve the eigenequation

$$\sum_{j'}^{A}(U_{jj'}^{A} - E\delta_{jj'})a_{j'}(\mathbf{k}) = 0, \qquad (2.28)$$

where

$$U_{jj'}^{A} = H_{jj'} + \sum_{\gamma\neq j,j'}^{B} \frac{H_{j\gamma}H_{\gamma j'}}{E_0 - E_\gamma} = H_{jj'} + \sum_{\gamma\neq j,j'}^{B} \frac{H'_{j\gamma}H'_{\gamma j'}}{E_0 - E_\gamma}, \qquad (2.29)$$

$$H'_{j\gamma} = \left\langle u_{j0}\left|\frac{\hbar}{m_0}\mathbf{k}\cdot(\mathbf{p} + \frac{\hbar}{4m_0c^2}\sigma\times\nabla V)\right|u_{\gamma0}\right\rangle \approx \sum_{\alpha}\frac{\hbar k_\alpha}{m_0}p_{j\gamma}^{\alpha}. \qquad (2.30)$$

As mentioned above, the second term in Π can be neglected compared to the similar term with $\mathbf{p}$ instead of $\mathbf{k}$. Similarly to the single band case, we can write for $U_{jj'}^A$

$$D_{jj'} \equiv U_{jj'}^A = E_j(0)\delta_{jj'} + \sum_{\alpha,\beta} D_{jj'}^{\alpha\beta} k_\alpha k_\beta, \qquad (2.31)$$

$$D_{jj'}^{\alpha\beta} = \frac{\hbar^2}{2m_0}\left[\delta_{jj'}\delta_{\alpha\beta} + \sum_\gamma^B \frac{p_{j\gamma}^\alpha p_{\gamma j'}^\beta + p_{j\gamma}^\beta p_{\gamma j'}^\alpha}{m_0(E_0 - E_\gamma)}\right]. \qquad (2.32)$$

We now define [19]

$$A_0 = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_\gamma^B \frac{p_{x\gamma}^x p_{\gamma x}^x}{E_0 - E_\gamma},$$

$$B_0 = \frac{\hbar^2}{2m_0} + \frac{\hbar^2}{m_0^2}\sum_\gamma^B \frac{p_{x\gamma}^y p_{\gamma x}^y}{E_0 - E_\gamma},$$

$$C_0 = \frac{\hbar^2}{m_0^2}\sum_\gamma^B \frac{p_{x\gamma}^x p_{\gamma y}^y + p_{x\gamma}^y p_{\gamma y}^x}{E_0 - E_\gamma},$$

and the band structure parameters (Luttinger parameters)

$$-\frac{\hbar^2}{2m_0}\gamma_1 = \frac{1}{3}(A_0 + 2B_0),$$

$$-\frac{\hbar^2}{2m_0}\gamma_2 = \frac{1}{6}(A_0 - B_0),$$

$$-\frac{\hbar^2}{2m_0}\gamma_3 = \frac{C_0}{6}.$$

These parameters are very closely related to the effective masses of the holes in the various valence bands. γ_1 and γ_2 describe the coupling of the $|X\rangle$, $|Y\rangle$ and $|Z\rangle$ states to the other states. The third parameter γ_3 relates to the anisotropy of the energy band structure around the Γ point when $\gamma_3 \neq \gamma_2$.

The Luttinger-Kohn Hamiltonian matrix $\mathbf{D}_{jj'}$ can be written explicitly as [18, 17, 20]:

$$
\mathbf{H} = \begin{bmatrix}
E_{el} & P_z & \sqrt{2}P_z & -\sqrt{3}P_+ & 0 & \sqrt{2}P_- & P_- & 0 \\
P_z^\dagger & P+\Delta & \sqrt{2}Q^\dagger & -S^\dagger/\sqrt{2} & -\sqrt{2}P_+^\dagger & 0 & -\sqrt{3/2}S & -\sqrt{2}R \\
\sqrt{2}P_z^\dagger & \sqrt{2}Q & P-Q & -S^\dagger & -P_+^\dagger & \sqrt{3/2}S & 0 & R \\
-\sqrt{3}P_+^\dagger & -S/\sqrt{2} & -S & P+Q & 0 & \sqrt{2}R & R & 0 \\
0 & -\sqrt{2}P_+ & -P_+ & 0 & E_{el} & P_z & -\sqrt{2}P_z & -\sqrt{3}P_- \\
\sqrt{2}P_-^\dagger & 0 & \sqrt{3/2}S^\dagger & \sqrt{2}R^\dagger & P_z^\dagger & P+\Delta & \sqrt{2}Q^\dagger & -S/\sqrt{2} \\
P_-^\dagger & -\sqrt{3/2}S^\dagger & 0 & R^\dagger & -\sqrt{2}P_z^\dagger & \sqrt{2}Q & P-Q & S \\
0 & -\sqrt{2}R^\dagger & R^\dagger & 0 & -\sqrt{3}P_-^\dagger & -S^\dagger/\sqrt{2} & S^\dagger & P+Q
\end{bmatrix}
$$

with [20]

$$
E_{el} = E_g + \frac{\hbar^2}{2m_e}(k_x^2 + k_y^2 + k_z^2), \tag{2.33}
$$

$$
P = \frac{\hbar^2\gamma_1}{2m_0}(k_x^2 + k_y^2 + k_z^2), \tag{2.34}
$$

$$
P_\pm = \sqrt{\frac{1}{6}}\left[i\mathcal{P}(k_x \pm ik_y) + \mathcal{B}k_z(k_y \pm ik_x)\right], \tag{2.35}
$$

$$
P_z = \sqrt{\frac{1}{3}}\left(i\mathcal{P}k_x + \mathcal{B}k_xk_y\right), \tag{2.36}
$$

$$
Q = \frac{\hbar^2\gamma_2}{2m_0}(k_x^2 + k_y^2 - 2k_z^2), \tag{2.37}
$$

$$
R = \frac{\hbar^2}{2m_0}[-\sqrt{3}\gamma_2(k_x^2 - k_y^2) + i2\sqrt{3}\gamma_3 k_x k_y], \tag{2.38}
$$

$$
S = \frac{\hbar^2\gamma_3}{m_0}\sqrt{3}(k_x - ik_y)k_z. \tag{2.39}
$$

Here Δ is the spin-orbit splitting energy. The coupling between the Γ conduction band-edge $|S\rangle$ state and the Γ valence-band-edge state $|Z\rangle$ is given by:

$$
\mathcal{P} = -\frac{\hbar^2}{m_0}\int_{unit\ cell} \psi_s \frac{\partial}{\partial z}\psi_z. \tag{2.40}
$$

The Kane parameter $\mathcal{B}$ describes the inversion asymmetry. In most practical calculations, this parameter is neglected [21]. The parameters γ_1, γ_2, γ_3 and $\mathcal{P}$ can be determined from effective masses at the Γ point of a bulk semiconductor [20]:

$$\frac{m_0}{m_{hh}(001)} = \gamma_1 - 2\gamma_2, \tag{2.41}$$

$$\frac{m_0}{m_{lh}(001)} = \gamma_1 + 2\gamma_2 + \lambda, \tag{2.42}$$

$$\frac{m_0}{m_{SO}(001)} = \gamma_1 + \frac{1}{2}\lambda r, \tag{2.43}$$

$$\frac{m_0}{m_{hh}(111)} = \gamma_1 - 2\gamma_3, \tag{2.44}$$

where the dimensionless parameters λ and r are given by

$$\lambda = \frac{4m_0\mathcal{P}^2}{3\hbar^2 E_g}, \tag{2.45}$$

$$r = \frac{E_g}{E_g + \Delta}. \tag{2.46}$$

Some values for the effective masses of GaAs and AlGaAs are given in table (2.1).

Material	m_{el}	m_{hh}	m_{lh}	m_{SO}	m_{hh}^{111}	source
GaAs	0.067	0.4537	0.0700	0.1434	0.8526	[20]
GaAs		0.465	0.085		0.595	cyclotron resonance [22]
GaAs		0.474	0.076		0.68	calculation [23] in [24]
GaAs	0.067	0.51	0.082	0.154	0.63	[25]
GaAs	0.067	0.62	0.087	0.15		[26]
Al$_{.25}$Ga$_{.75}$As	0.0942	0.5100	0.0900	0.1720	0.9815	[20]
AlAs	0.150	0.76	0.15	0.24		[26]

Table 2.1: *Effective mass at the valence band Γ point, according to different sources*

2.3.3   2 Band Model

The conduction band can be modeled quite easily if we assume that the interaction with the other bands is weak enough for it to be treated perturbatively, i.e. use a simple effective mass model. In the case of the valence band, however, the strong interaction between the degenerate light and heavy hole bands (near the band edge) requires that these bands are taken into account explicitly. Only when we consider energy levels deep into the valence bands (close to the SO splitting energy, about 300 meV in GaAs) do the coupling terms to the SO band become important. As we will only be concerned with shallow levels, the influence of SO and conduction bands (1.5 eV splitting) can be introduced through the effective mass.

The degeneracy of the light and heavy hole bands near the band edge generates a coupling term (as in the Luttinger-Kohn Hamiltonian). Including spin degeneracy, this yields a set of four coupled effective mass equations [17, 21, 27].

Fortunately, this set of coupled equations can be greatly simplified by a method described by Broido and Sham [21]. They used a unitary transformation of the four basis Bloch functions $(u_{lh}, u_{hh}, \bar{u}_{lh}, \bar{u}_{hh})$ into a new set (u_A, u_B, u_C, u_D) to decouple the set of four coupled equations into two sets of two coupled equations. The Bloch functions u_i are given by

$$u_A = \frac{1}{\sqrt{2}}(u_{hh} - \bar{u}_{hh}),$$
(2.47)

$$u_B = \frac{1}{\sqrt{2}}(-u_{lh} + \bar{u}_{lh}),$$
(2.48)

$$u_C = \frac{1}{\sqrt{2}}(u_{lh} + \bar{u}_{lh}),$$
(2.49)

$$u_D = \frac{1}{\sqrt{2}}(u_{hh} + \bar{u}_{hh}).$$
(2.50)

This means that the Hamiltonian

$$
H = \begin{bmatrix}
P - Q & -S^\dagger & 0 & R \\
-S & P + Q & R & 0 \\
0 & R^\dagger & P - Q & S \\
R^\dagger & 0 & S^\dagger & P + Q
\end{bmatrix}
\tag{2.51}
$$

is diagonalized into two 2×2 blocks, an upper block H^U and lower block H^L, given by

$$
H^\sigma = \begin{bmatrix}
P \pm Q & W \\
W^\dagger & P \mp Q
\end{bmatrix}, W = |R| - i|S|,
\tag{2.52}
$$

where the index $\sigma = U(L)$ refers to the upper (lower) $\pm$ signs. The upper and lower blocks are equivalent, showing the double degeneracy of the heavy and light hole bands. It is therefore sufficient to solve the upper block and obtain its solutions. The solutions for the lower block can easily be determined from the latter.

We can identify $(P - Q)$ and $(P + Q)$ with the light hole energy (operator) H_{lh} and the heavy hole energy H_{hh}, respectively. Similarly to the conduction band case, the Schrödinger equation with Hamiltonian (2.52) can be simplified into an effective-mass formalism with:

$$
\hat{H}_{lh} = -(\gamma_1 + 2\gamma_2)\frac{\partial^2}{\partial z^2} + (\gamma_1 - \gamma_2)k_t^2,
\tag{2.53}
$$

$$
\hat{H}_{hh} = -(\gamma_1 - 2\gamma_2)\frac{\partial^2}{\partial z^2} + (\gamma_1 + \gamma_2)k_t^2,
\tag{2.54}
$$

$$
\hat{W} = \sqrt{3}k_t(\gamma_2 k_t - 2\gamma_3\frac{\partial}{\partial z}).
\tag{2.55}
$$

Finally, we take into account a potential $V(z)$, which represents the (bulk) valence-band-edge offset with respect to an arbitrary reference energy. This allows us to write the effective mass equation as :

$$
\begin{bmatrix}
H_{hh} + V & W \\
W^\dagger & H_{lh} + V
\end{bmatrix}
\begin{bmatrix}
F_{hh} \\
F_{lh}
\end{bmatrix}
= E(\mathbf{k})
\begin{bmatrix}
F_{hh} \\
F_{lh}
\end{bmatrix},
\tag{2.56}
$$
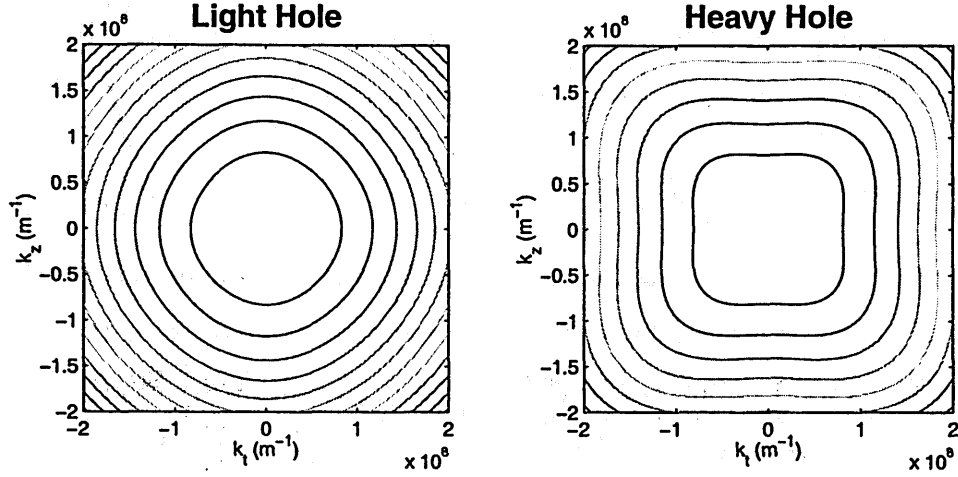
Figure 2-2: *Constant energy contours for light holes and heavy holes in bulk GaAs. There is a clear anisotropy in the* ⟨110⟩ *direction. The energy spacing is 0.5 meV for the HH band and 3 meV for the LH band.*

where F_{hh} and F_{lh} are the envelope functions corresponding to u_A and u_B respectively. Note that in this formalism, hole energies are taken to be positive.

The first step in solving the quantum well problem, is finding the solution in bulk material, where we take V to be a constant V_0. The value of V_0 will be different in well material and barriers, reflecting the different valence band edge offsets. We can now easily solve for the eigenenergies $E(\mathbf{k})$, yielding the bulk energy dispersion relations for the HH and LH bands. We consider the case of a {100} plane, writing the in-plane $\mathbf{k}$ component as k_t :

$$E(\mathbf{k}) - V_0 = \gamma_1(k_z^2 + k_t^2) \pm \sqrt{4\gamma_2^2(k_z^2 + k_t^2)^2 + 12(\gamma_3^2 - \gamma_2^2)k_z^2 k_t^2}, \qquad (2.57)$$

where the plus sign refers to the "light hole" solution, and the minus sign to the "heavy hole" solution. We can rewrite (2.57) as:

$$E(\mathbf{k}) - V_0 = \left[\gamma_1 \pm \gamma_2\sqrt{1 + 3\frac{\gamma_3^2 - \gamma_2^2}{\gamma_2^2}\frac{k_z^2 k_t^2}{(k_z^2 + k_t^2)^2}}\right](k_z^2 + k_t^2). \qquad (2.58)$$

Constant energy contours are shown in figure (2-2), illustrating that γ_3 can be related to the mass anisotropy along the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions. If k_t is small compared to k_z, we can expand the square root in (2.57):

$$E(\mathbf{k}) - V_0 = (\gamma_1 \pm \gamma_2)(k_z^2 + k_t^2) \pm 3\frac{\gamma_3^2 - \gamma_2^2}{\gamma_2}k_t^2. \tag{2.59}$$

The energy term accounting for anisotropy for a given k_t and k_z is equal for the HH and LH bands. However, due to the lower energy of the HH bands the anisotropy term is relatively more important for HH than for LH, resulting in a clearly anisotropic HH band and a quasi isotropic LH band.

Still, we see that in bulk material, the effective masses along the z-axis $\langle 001 \rangle$ and x- and y- axes $\langle 100 \rangle$ and $\langle 010 \rangle$ are identical (as expected), as the dispersion relation is given by $E(\mathbf{k}) - V_0 = (\gamma_1 \pm 2\gamma_2)k^2$. We can easily find this from (2.57) with $k_t = 0$ for $\langle 001 \rangle$, and $k_z = 0$ for the x- and y- directions.

The eigenvectors of (2.56) are found to be, apart from a normalization constant :

$$\psi_1(\mathbf{k}, \mathbf{r}) = \begin{bmatrix} F_{hh,1} \\ F_{lh,1} \end{bmatrix} = e^{i\mathbf{k}\cdot\mathbf{r}} \begin{bmatrix} H_{lh} + V_0 - E_{hh} \\ -W^{\dagger} \end{bmatrix}, \tag{2.60}$$

$$\psi_2(\mathbf{k}, \mathbf{r}) = \begin{bmatrix} F_{hh,2} \\ F_{lh,2} \end{bmatrix} = e^{i\mathbf{k}\cdot\mathbf{r}} \begin{bmatrix} H_{lh} + V_0 - E_{lh} \\ -W^{\dagger} \end{bmatrix}, \tag{2.61}$$

where the matrix notation implies

$$\psi = F_{hh}u_A + F_{lh}u_B. \tag{2.62}$$

To solve the quantum well problem, we choose the well growth direction (direction of confinement) along the z-axis. The xy-plane is in the plane of the well. We can construct a confined solution from the bulk plane wave solutions by imposing boundary conditions along the confinement axis. In the plane of the well, there is no confinement and hence we retain the bulk plane wave solutions. By taking a linear combination of the bulk solutions in each material, a general solution can be
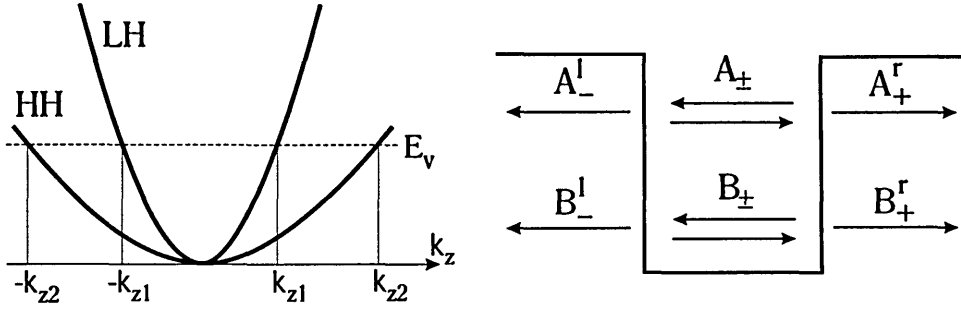
Figure 2-3: *At any one energy in a bulk material, we can find four wavevectors corresponding to the heavy and light hole bands. An eigenstate of the Hamiltonian in a quantum well is then made up of a linear combination of the bulk plane waves corresponding to those wave vectors. $A_\pm$ corresponds to $\pm k_{z1}$ (light hole), $B_\pm$ to the heavy hole $\pm k_{z2}$ in GaAs. In the barriers (AlGaAs) a similar mechanism is employed. The boundary conditions at the interfaces then determine the energy eigenvalues and the coefficients. Only the outgoing waves are indicated, as the coefficients for the incoming waves have to be identically zero.*

constructed. As illustrated in figure (2-3), four plane wave solutions exist at a given energy, yielding a general solution Ψ:

$$\Psi = \sum A_\pm \psi_1(\pm k_{z1}, k_t, \mathbf{r}) + \sum B_\pm \psi_2(\pm k_{z2}, k_t, \mathbf{r}). \tag{2.63}$$

The four coefficients $A_\pm$ and $B_\pm$ are unknown constants. Both ψ_1 and ψ_2 are two-component vectors, as described in (2.60) and (2.61). We can write the components of Ψ, F_{hh} and F_{lh}, as :

$$F_{hh} = e^{ik_t \cdot r_t} \left[\sum A_\pm F_{hh,1}(\pm k_{z1}, k_t) e^{\pm i k_{hh} z} + \sum B_\pm F_{hh,2}(\pm k_{z2}, k_t) e^{\pm i k_{hh} z} \right], \tag{2.64}$$

$$F_{lh} = e^{ik_t \cdot r_t} \left[\sum A_\pm F_{lh,1}(\pm k_{z1}, k_t) e^{\pm i k_{lh} z} + \sum B_\pm F_{lh,2}(\pm k_{z2}, k_t) e^{\pm i k_{lh} z} \right]. \tag{2.65}$$

Thus we have four unknown constants in each region, making a total of twelve unknowns over the three regions. The boundary conditions at the interfaces between the regions and the demand that the solutions be confined in the quantum well provide the necessary relations to solve the problem. The following quantities have

Valence Band Energy vs k for a 100 A GaAs/AlGaAs QW
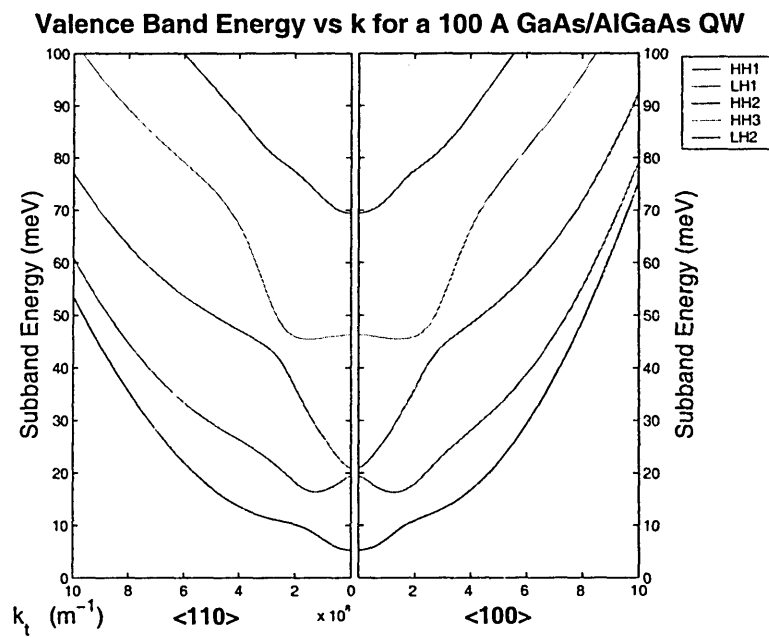
Figure 2-4: *Dispersion relations for $\langle 100 \rangle$ and $\langle 110 \rangle$ directions in a 100 Å wide $GaAs/Al_{.3}Ga_{.7}As$ quantum well. Growth direction is $\langle 001 \rangle$. The subbands are named after their dominant character at the zone center. At higher k-vectors the bands are very heavily mixed.*
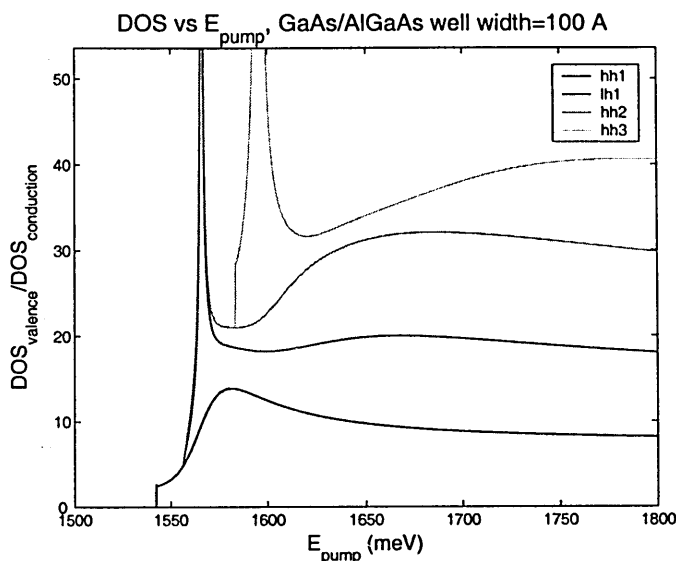
Figure 2-5: *Density of states in a 100 Å wide GaAs/Al.3Ga.7As quantum well. The spikes in the density of states for LH1 and HH3 are due to band extrema away from the zone center.*

to be matched across an interface :

$$F_{hh} \quad \text{and} \quad (\gamma_1 - 2\gamma_2)\frac{dF_{hh}}{dz} + \sqrt{3}\gamma_3 k_t F_{lh}, \tag{2.66}$$

$$F_{lh} \quad \text{and} \quad (\gamma_1 - 2\gamma_2)\frac{dF_{lh}}{dz} - \sqrt{3}\gamma_3 k_t F_{hh}. \tag{2.67}$$

These boundary conditions were obtained by symmetrizing the Hamiltonian in (2.52). Caution should be issued however that the above boundary conditions only apply when the Bloch functions in both well materials are similar (also as seen from their Luttinger parameters), as is the case for the GaAs-AlGaAs system. The boundary conditions boil down to the continuity of the wave function and "generalized" continuity of its derivative, corresponding to current across the interface.

In figure (2-4), the valence subband structure of a 100Å GaAs/Al$_{0.3}$Ga$_{0.7}$As quantum well is shown ($V_0 \approx 140$ meV if the reference is the GaAs valence band edge). The LH and HH bands are very heavily coupled, giving rise to highly non-parabolic subbands.

Also notice the anisotropy between $\langle 100 \rangle$ and $\langle 110 \rangle$ directions. In practical calculations, we can use an approximation where the coupling terms which are responsible for the anisotropy are replaced by an "average" term $W_{ave} = (W_{100} + W_{110})/2$. This approach is known as the axial approximation.

Particularly important is the density of states of the subbands, as illustrated in figure (2-5), as this plays a major role in determining interband gain and absorption. The density of states $g(E)$ can be found from

$$g(E) = \frac{1}{\pi} \frac{dk}{dE},\tag{2.68}$$

assuming the E-k relationship is isotropic (using the axial approximation). The DOS starts off at roughly 2.5 times the conduction band DOS ($2.8 \times 10^{10} \text{cm}^{-2}\text{meV}^{-1}$), increasing with increasing hole energy in the valence band.

2.4   Optical Transitions

At the core of an optical device, is of course the interaction of electromagnetic waves (light) and matter (electrons in the semiconductor). Quantum mechanically, the interaction between photons and electrons in the semiconductor can be described by the Hamiltonian

$$H = \frac{1}{2m_0}(\mathbf{p} - q\mathbf{A})^2 + V(\mathbf{r}),\tag{2.69}$$

where $\mathbf{A}$ is the magnetic vector potential, and q is the carrier charge ($q = -e$ for electrons). Neglecting the term quadratic in $\mathbf{A}$ (a good approximation for most practical optical field intensities), and applying the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$, we can distinguish the perturbation Hamiltonian H' due to the electron-photon interaction:

$$H' \approx \frac{e}{m_0}\mathbf{A} \cdot \mathbf{p}.\tag{2.70}$$

However, it is possible to take the quadratic term into account without making the perturbation Hamiltonian more cumbersome. We can do this by by explicitly writing

the wavefunction ψ as the product of a phase factor and the remaining wavefunction ψ':

$$\psi \to \psi' \, e^{ie\mathbf{A}\cdot\mathbf{r}/\hbar}. \tag{2.71}$$

Substituting this into the Schrödinger equation with Hamiltonian (2.69), we find:

$$\mathbf{p}\psi = e^{ie\mathbf{A}\cdot\mathbf{r}/\hbar}\,(\mathbf{p} + e\mathbf{A})\psi', \tag{2.72}$$

$$\frac{p^2}{2m_0}\psi' = \left[\frac{\partial}{\partial t} - e\frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{r}\right]\psi', \tag{2.73}$$

and, as $\mathbf{E} = -\partial\mathbf{A}/\partial t$, the Schrödinger equation for the interaction of a photon with an electron can be written as a function of the amplitude of the incident optical field $\mathbf{E}$:

$$\left[\frac{p^2}{2m_0} - e\mathbf{E}\cdot\mathbf{r}\right]\psi' = i\hbar\frac{\partial\psi'}{\partial t}, \tag{2.74}$$

The perturbation Hamiltonian is:

$$H' = -e\mathbf{E}\cdot\mathbf{r}. \tag{2.75}$$

The physical interpretation of this interaction is more intuitively obvious than in the description with a vector potential. The radiative field acts as a force on the electron charge cloud, thus accelerating it and generating radiation (emitting a photon) or exciting the electron (absorption of a photon).

In order to describe the particle-particle interaction between photon and electron, we have to quantize the electric field E. This can be done similarly to the case of a harmonic oscillator. A photon then corresponds to one quantum of excitation in an oscillator:

$$E(\mathbf{r}, t) = -i\sqrt{\frac{\hbar\omega}{2V\epsilon}}\left[\mathbf{a}^\dagger e^{-i\kappa\cdot\mathbf{r}+i\omega t} - \mathbf{a}e^{i\kappa\cdot\mathbf{r}-i\omega t}\right]. \tag{2.76}$$

The operators $\mathbf{a}$ and $\mathbf{a}^\dagger$ are photon annihilation and creation operators, respectively. They correspond to the absorption and emission of photons by mediation of

an oscillating electric (electromagnetic) field with angular frequency ω. This is true even if that field is a vacuum field, as is the case for spontaneous emission.

The optical transition rate between an initial state $(E_i, \mathbf{k_i})$ and final state $(E_f, \mathbf{k_f})$ can be described using Fermi's Golden Rule:

$$W_{if} = \frac{2\pi}{\hbar} \left| \langle \psi_f | H'(\mathbf{r}) | \psi_i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega). \tag{2.77}$$

Here the delta function assumes a zero linewidth. In order to introduce a finite linewidth, the delta function can be replaced with the proper line-shape, usually a Lorentzian with linewidth Γ. The Lorentzian is a good model for line-shape broadening due to a finite lifetime or dephasing scattering.

$$\delta(E_f - E_i - \hbar\omega) \rightarrow \frac{\Gamma/(2\pi)}{(E_f - E_i - \hbar\omega)^2 + (\Gamma/2)^2}. \tag{2.78}$$

More generally, in most cases a number of final states is available, with density of states $\rho(E_{if})$. As each state is equally probable as a final state for the transition, we obtain (zero linewidth)

$$W_{if} = \frac{2\pi}{\hbar} \left| \langle \psi_f | H'(\mathbf{r}) | \psi_i \rangle \right|^2 \rho(E_{if})\delta(E_f - E_i - \hbar\omega). \tag{2.79}$$

If we neglect non-parabolicity, the subbands in one particular band track each other. The energy separation between two states with identical in-plane wave vector remains constant for any two given subbands. Assuming only vertical transitions (dipole selection rule), this means that ρ will be given by the subband density of states for intersubband transitions. Neglecting non-parabolicity is okay for the conduction band, but not for the valence band. Band-mixing in valence band quantum wells leads to large non-parabolicity effects.

In interband transitions, the bands don't track each other and the energy separation varies with E_{if}. Again, assuming only vertical transitions in k-space, the number of transition pairs within δk has to be the same in both conduction band and valence band $\rho_r \, \delta E_{cv} = \rho_c \, \delta E_c = \rho_v \, \delta E_v$. Setting $\delta E_{cv} = \delta E_c + \delta E_v$, we immediately obtain:

$$\frac{1}{\rho_r} = \frac{1}{\rho_c} + \frac{1}{\rho_v}. \tag{2.80}$$
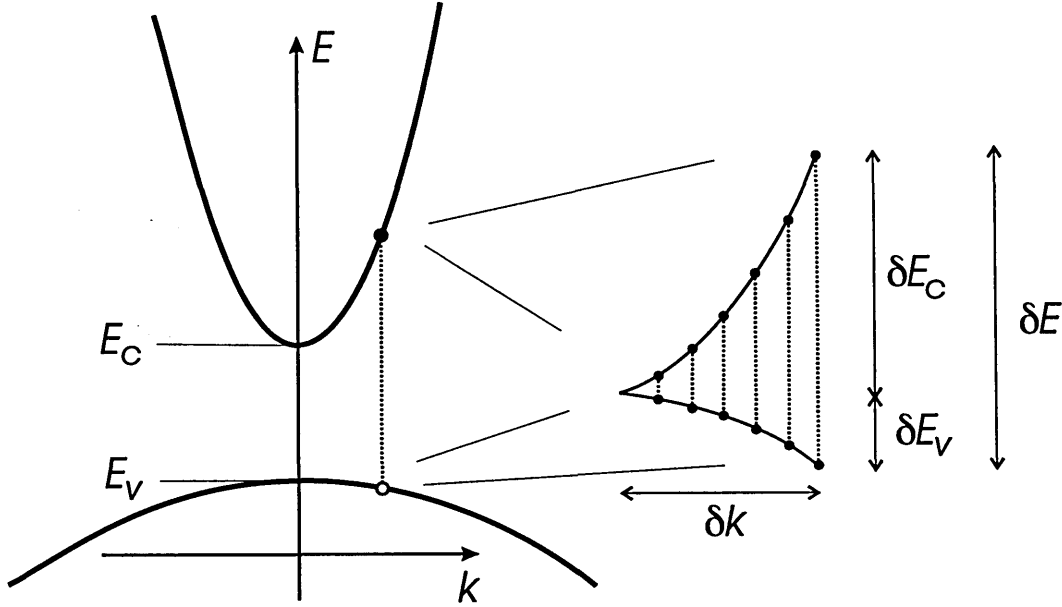
Figure 2-6: *Relationship between the energy ranges in the conduction and valence bands for a given dk in k-space, assuming k-selection applies.*

In typical semiconductors, like GaAs, ρ_v is a lot higher than ρ_c, hence $\rho_r \approx \rho_c$. The above equation can be rewritten in a more practical form, useful for non-parabolic bands:

$$\frac{1}{\rho(E_{if})} = \frac{1}{\rho(k)} \left[\frac{dE_c(k)}{dk} - \frac{dE_v(k)}{dk} \right]. \tag{2.81}$$

This definition allows for ρ_r to be evaluated at any given point in k-space once the $E - k$ dispersion relation is known at that point.

2.4.1   Intersubband Optical Transitions

We can write the initial (photon density n_{ph}) and final states $(n_{ph} + 1)$ in an intersubband transition as:

$$\psi_i = \frac{e^{i\mathbf{k}_{t,i}\cdot\mathbf{r}_t}}{\sqrt{A}} F_i(z) u_{0,i}, \tag{2.82}$$

$$\psi_f = \frac{e^{i\mathbf{k}_{t,f} \cdot \mathbf{r}_t}}{\sqrt{A}} F_f(z) u_{0,f}, \tag{2.83}$$

where A is the in-plane area of the quantum well. As both the Bloch functions u belong to the same band, and as they are almost independent of $\mathbf{k}$, we can assume $u_{0,i} \approx u_{0,f}$. The matrix element H_{if} can then be written:

$$H_{if} \approx -i\sqrt{\frac{\hbar\omega(n_{ph}+1)}{2V\epsilon}} \; \hat{\mathbf{e}} \cdot \left\langle \frac{e^{i\mathbf{k}_{t,f} \cdot \mathbf{r}_t}}{\sqrt{A}} F_f(z)|e\mathbf{r}|\frac{e^{i\mathbf{k}_{t,i} \cdot \mathbf{r}_t}}{\sqrt{A}} F_i(z) \right\rangle \; \langle u_{0,f}|u_{0,i}\rangle, \tag{2.84}$$

$$H_{if} \approx -i\sqrt{\frac{\hbar\omega(n_{ph}+1)}{2V\epsilon}} \; \hat{\mathbf{e}} \cdot \langle F_f(z)|e\mathbf{z}|F_i(z)\rangle \; \delta_{\mathbf{k}_{t,f},\mathbf{k}_{t,i}}. \tag{2.85}$$

In the above equation we made use of the fact that the $F(z)$ can be considered a constant on the scale of a lattice spacing. This is a very good assumption for low level states and quantum wells wider than a few monolayers. This is generally true for the wave functions we are interested in.

The delta function in (2.85) corresponds to a conservation of in-plane momentum. The momentum carried by the photon, $\mathbf{k}_{ph}$, is of the order of $2\pi/\lambda$ ($\lambda \approx 100\mu m$), which is negligible compared to the electron wave vectors $\mathbf{k}_i$, $\mathbf{k}_f$ of the order of $2\pi/a$, a being the lattice constant (order of magnitude 5 Å). Therefore we can write $\mathbf{k}_{t,f} \approx \mathbf{k}_{t,i}$.

The matrix element $Z_{if} = \langle F_f|\mathbf{z}|F_i\rangle$ is called the dipole matrix element between the initial and final states. The dipole matrix element can be used as a gauge for the strength of the optical intersubband transition. Due to the dimensions of the quantum wells and their (bound) energy levels, Z_{if} in intersubband transitions (~ 30 Å) can be a lot larger than in an atomic system (~ 2 Å). During the design of a quantum well structure, we will try to maximize the dipole moment associated with the targeted intersubband transition.

Also apparent from (2.85) is a dipole selection rule for intersubband transitions. Only an electromagnetic wave with its electric field polarized along the z-axis (the quantum well growth axis) will be generated or absorbed in an intersubband transition.

Using Fermi's Golden Rule, the intersubband transition rate for stimulated emission (into one specific optical mode, i.e. the same one as the incident wave) can be written as:

$$W_{if} = \frac{\pi e^2 \omega Z_{if}^2 n_{ph}}{V \epsilon} \delta(E_f - E_i - \hbar\omega).$$

(2.86)

The transition rate is directly proportional to the intensity of the incident field ($\sim n_{ph}$). Equation 2.86 also shows that the transition rate decreases with increasing wavelength. The expression for (stimulated) absorption is identical to the one for stimulated emission.

For spontaneous intersubband emission, we have to sum over all available final photon states. Taking into account a 3D optical mode density of $(8\pi V n^3 E^2)/(h^3 c^3)$, the transition rate is:

$$W_{if,sp}^{3D} = \frac{e^2 n \omega^3 Z_{if}^2}{3\pi \epsilon_0 \hbar c^3}.$$

(2.87)

However, in far-infrared optical quantum electronic devices, the transition usually takes place inside a two-dimensional optical cavity with thickness t_c which is at the same scale or smaller than the wavelength (50-100 μm). This cavity can consist of a metal or plasma waveguide, confining the electromagnetic wave in the z-direction and limiting the optical mode density to $A/(2\pi)^2$. This yields a 2D intersubband transition rate of

$$W_{if,sp}^{2D} = \frac{e^2 n \omega^2 Z_{if}^2}{2 t_c \epsilon_0 \hbar c^2},$$

(2.88)

scaling inversely proportional to the cavity thickness. Compared to the 3D expression, this dependence replaces a $1/\lambda$ dependence. This is shown more clearly if we look at the ratio of W^{3D} to W^{2D}

$$\frac{W_{if,sp}^{3D}}{W_{if,sp}^{2D}} = \frac{4 t_c}{3\lambda}.$$

(2.89)

The microcavity effect will increase W^{2D} over the 3D case if the thickness of the cavity is smaller than the wavelength. Note that a microcavity only has an effect on

the spontaneous emission rate. Stimulated emission, and hence gain, are not affected as all photons are coupled into one single mode. How many modes are available, is not important.

Optical gain is defined as the relative increase of a wave intensity per length unit as the wave propagates through the medium : $dI/dx = g(\omega)I$. To find the expression for optical gain, we subtract total absorption $\hbar\omega N_f W_{ab}$ from total stimulated emission $\hbar\omega N_i W_{st}$. With pump beam intensity $I = \frac{n_{ph}}{V}\hbar\omega\frac{c}{n}$ we find from (2.86):

$$W_{st} = \frac{e^2 n Z_{ij}^2}{2\hbar^2 \epsilon c} I\delta(E_f - E_i - \hbar\omega),\tag{2.90}$$

and

$$g(\omega) = \frac{\Delta N e^2 \omega Z_{ij}^2}{2\hbar n \epsilon_0 c}\delta(E_f - E_i - \hbar\omega).\tag{2.91}$$

Here $\Delta N = N_i - N_f$ is the population inversion between initial and final subbands. If the transition has a finite linewidth Δf, the delta function in (2.86) is replaced with a Lorentzian line-shape and the maximum gain is:

$$g_0 = \frac{\Delta N e^2 \omega Z_{ij}^2}{\pi \hbar^2 n \epsilon_0 c \Delta f}.\tag{2.92}$$

2.4.2   Interband Optical Transitions

The initial and final states in an interband transition between conduction band and valence band are

$$\psi_i = \frac{e^{i\mathbf{k_{t,i}}\cdot\mathbf{r_t}}}{\sqrt{A}}F_i(z)u_c,\tag{2.93}$$

$$\psi_f = \frac{e^{i\mathbf{k_{t,f}}\cdot\mathbf{r_t}}}{\sqrt{A}}F_f(z)u_v.\tag{2.94}$$

The optical interband matrix element H_{if} now takes a slightly different form than for the intersubband case.

$$H_{if} = \frac{qA}{2m_0}\langle u_c|\hat{e}\cdot\mathbf{p}|u_v\rangle\ \delta_{\mathbf{k_{t,f}k_{t,i}}}\ \langle F_f|F_i\rangle.\tag{2.95}$$
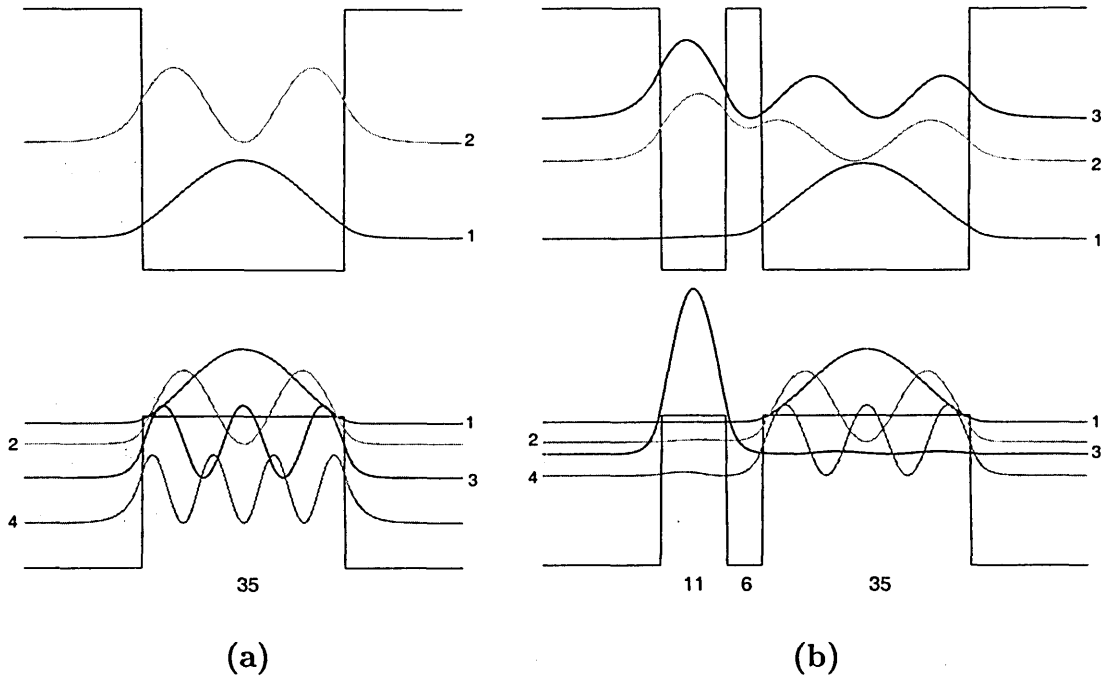
Figure 2-7: (a) *Subband edge probability distributions for the conduction and heavy hole valence bands in a single 100 Å $GaAs/Al_{.3}Ga_{.7}As$ QW. Growth direction is $\langle 001 \rangle$.* (b) *Subband edge wavefunctions for the conduction and heavy hole valence bands in a pair of coupled $GaAs/Al_{.3}Ga_{.7}As$ QW's. Only the four lowest energy heavy hole subbands are illustrated. The width of barriers and wells is indicated in monolayers (1ML=2.825Å). Growth direction is $\langle 001 \rangle$.*

Here we made use of (2.70) rather than (2.75). Notable is that the same k selection applies as in the intersubband case. The vector ê is the unit vector pointing along **A**. As $\mathbf{E} = -\partial \mathbf{A}/\partial t$, ê will also be parallel to **E** if the exciting field is linearly polarized.

The overlap integral $\langle F_f | F_i \rangle$ between the conduction and valence band envelope functions gives rise to another, less strict selection rule. Usually the overlap between states with the same quantum number in the same well is much higher than with other states. So, transitions between corresponding states in conduction and valence band will be favored ("allowed"), whereas the other transitions are less strong. However, the difference in effective mass between the bands and wave mixing can cause important exceptions to this rule. This is illustrated in figure (2-7) and table (2.2). For a single

well the subband edge wavefunctions would be perfectly orthogonal if they had the same effective mass and if the wells had the same depth. These differences between the valence and conduction band give rise to small violations of the selection rule, becoming more important as the involved subbands are higher in energy and thus closer to the barrier energy. In the double quantum well system the ground level of the narrow well and the first excited level of the wide well are heavily coupled. The four involved states ($\psi_{c2}, \psi_{c,3}, \psi_{v,2}, \psi_{v,3}$) now share a significant overlap. The overlap with the "non-perturbed" states ($\psi_{c,1}$ and $\psi_{v,1}, \psi_{v,4}$ in the example) remains small. Of course, the orthogonality of states in the same band (valence or conduction) still holds and the overlap integral between two conduction or valence band states is identically zero.

Cond \ Val	$\psi_{v,1}$	$\psi_{v,2}$	$\psi_{v,3}$	$\psi_{v,4}$
$\psi_{c,1}$	0.9810	0.0000	0.0102	0.0000
$\psi_{c,2}$	0.0000	0.9260	0.0000	0.0388

Cond \ Val	$\psi_{v,1}$	$\psi_{v,2}$	$\psi_{v,3}$	$\psi_{v,4}$
$\psi_{c,1}$	0.9714	0.0007	0.0109	0.0088
$\psi_{c,2}$	0.0075	0.5963	0.3361	0.0001
$\psi_{c,3}$	0.0071	0.3410	0.5847	0.0055

Table 2.2: *The square of the overlap integrals $|\langle F_f|F_i\rangle|^2$ between the various subband edge wavefunctions as illustrated in figure (2-7). The top table refers to the single well in (a), the bottom table refers to the coupled wells in (b).*

The momentum matrix element $\mathbf{M_T} = \langle u_c|\hat{\mathbf{e}} \cdot \mathbf{p}|u_v\rangle \langle F_c|F_v\rangle$ is polarization dependent [28]. When calculating M_T, all possible transitions between the involved conduction and valence bands (HH or LH) have to be taken into account, yielding:

$$|M_T|^2 = \frac{1}{2} \sum_{u_c,\bar{u}_c} \sum_{u_v,\bar{u}_v} |\langle u_c|\hat{\mathbf{e}} \cdot \mathbf{p}|u_v\rangle \langle F_c|F_v\rangle|^2 .$$

(2.96)
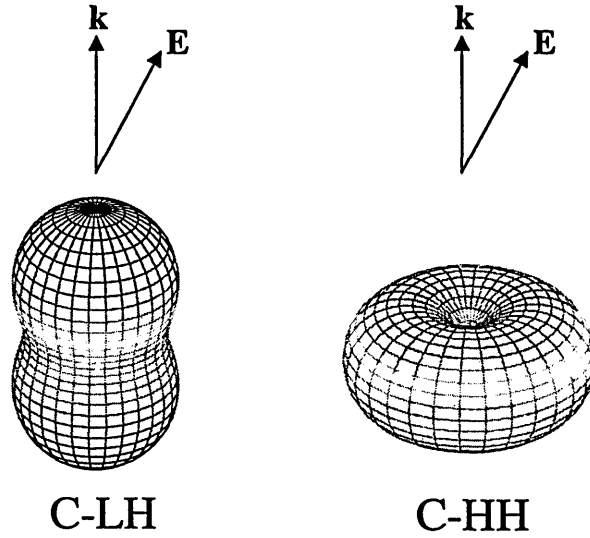
Figure 2-8: *Dependence of the transition strength $|M_T|$ on the angle between the electron k-vector and the the electric field polarization vector E.*

The factor 1/2 compensates for the spin degeneracy factor already present in the expression for the density of states. It is customary to express $|M_T|$ in function of $|M| \doteq |\langle u_c|\hat{e} \cdot \mathbf{p}|u_i\rangle|$, where $i = x, y, z$. $|M|$ can be determined from measurements of the band curvature [28]. By expressing the valence bands u_v as linear combinations of the basis functions u_i (equations (2.47)-(2.50) and (2.27)), the expression for $|M_T|^2$ can be much simplified. We assume the envelope function overlap integral to be unity, because for the moment we will be interested in transitions between two bulk plane wave states. We find for the normalized transition strength in bulk material [28]:

$$|M_T|^2/|M|^2 = \begin{cases} \frac{1}{2}(1 - |\hat{\mathbf{k}} \cdot \hat{e}|^2) & \text{for HH band,} \\ \frac{1}{2}(\frac{1}{3} + |\hat{\mathbf{k}} \cdot \hat{e}|^2) & \text{for LH band.} \end{cases} \qquad (2.97)$$

Here $\hat{\mathbf{k}}$ is a unit vector pointing along the electron k-vector.

As shown in figure (2-8), the strength of interaction between each electron plane wave state and an incident photon is highly polarization dependent. However, in bulk material this dependence doesn't reveal itself as the incident field interacts with

many electrons with different wave vectors, effectively averaging out the polarization difference. In bulk, the relative interaction strength is equal in all three principal directions $|M_T|^2/|M|^2 = 1/3$.

For TM polarized light ($\mathbf{E} \parallel \hat{z}$, $H_z = 0$, assuming propagation in the $\hat{z}$-direction), the heavy-hole transition is forbidden. This is due to the absence of a u_z component in the heavy-hole basis state u_{hh}. A p_z-like (odd in z) state is needed for the interaction of a z-polarized field with the s-like conduction band Bloch state.

In quantum well structures, we can no longer work with simple plane wave states, but have to take into account the envelope functions. As a general valence band wave-function consists of both HH and LH components, equation (2.96) can be replaced by

$$|M_T|^2 = \sum_{u_c, \bar{u}_c} |\langle u_c | \hat{e} \cdot \mathbf{p} | u_A \rangle \langle F_c | F_{hh} \rangle + \langle u_c | \hat{e} \cdot \mathbf{p} | u_B \rangle \langle F_c | F_{lh} \rangle|^2. \qquad (2.98)$$

Averaging out over all in-plane $\mathbf{k}$ directions, we can remove the cross-terms in the above equation and get explicit expressions for TE ($E_z = 0$, for beam propagation in z-direction) and TM ($H_z = 0$) light polarization, assuming the growth direction is $\hat{z}$:

$$|M_T|^2 = |M|^2 \left[\tfrac{2}{3} |\langle F_c | F_{lh} \rangle|^2 \right] \qquad \text{TM } (\hat{e} \parallel \hat{z}),$$

$$|M_T|^2 = \tfrac{|M|^2}{2} \left[|\langle F_c | F_{hh} \rangle|^2 + \tfrac{1}{3} |\langle F_c | F_{lh} \rangle|^2 \right] \qquad \text{TE } (\hat{e} \perp \hat{z}).$$

For the interband pumping scheme, TM polarization corresponds to edge pumping, while TE polarization relates to surface pumping.

At the band edge, the valence band states can be be characterized as pure HH or LH states. Apart from the overlap integral $|\langle F_f | F_i \rangle|^2$, this case corresponds to the bulk case and the obtained values for the LH and HH transition strength correspond to what was shown in figure (2-8). As k_t increases, band mixing becomes more and more important, drastically altering the transition strength. Away from the zone center, the states generally exhibit both HH and LH components. A sample calculation of a 100 Å quantum well is included in figure (2-9). For the heavy hole transition, absorption of TM polarized light is suppressed close to $k_t = 0$, where band mixing is
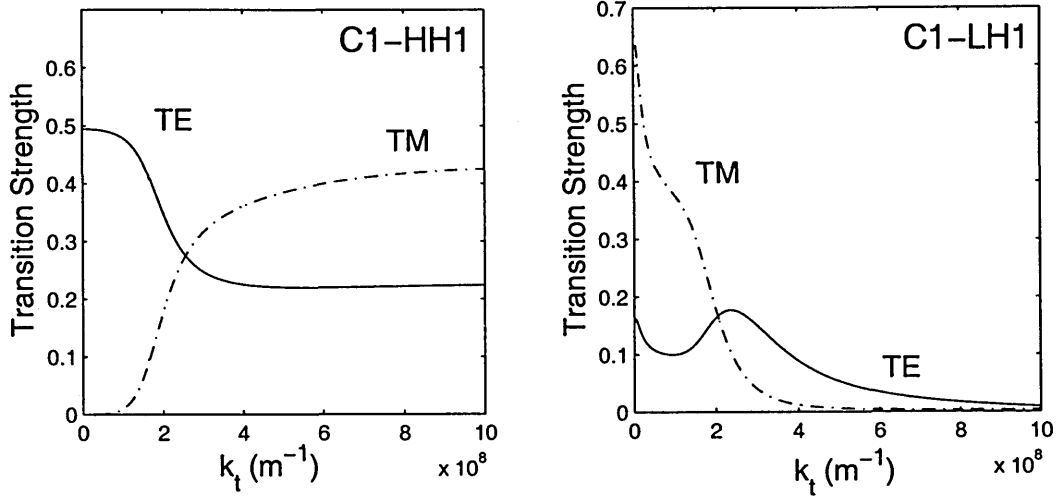
Figure 2-9: *Relative transition strengths for both TE and TM light polarization for the two lowest subband transitions in a 100 Å GaAs/Al.3Ga.7As quantum well.*

minimal. As we move away from the band edge, band mixing occurs such that both F_{lh} and F_{hh} are present in any one wavefunction. This alters the transition strengths.

We can use Fermi's Golden Rule to get an expression for the interband optical gain. The gain can be written as:

$$g_{if} = \frac{2\pi}{\hbar} \frac{|H_{if}|^2}{v_g n_{ph}} \rho_r(E_{if})(f_f - f_i),$$
(2.99)

with $v_g = c/n$ the group velocity of the incident wave. Substituting (2.95) in (2.99) yields the material gain per unit length:

$$g_{if} = \frac{q^2 \hbar}{\epsilon_0 c m_0^2} \frac{1}{\hbar \omega_{if}} |M_T(E_{if})|^2 \rho_r(E_{if})(f_f - f_i).$$
(2.100)

Chapter 3

Phonon-Carrier and Carrier-Carrier Scattering

3.1   Phonons

The atoms in a semiconductor lattice are linked together with chemical bonds. These bonds can be strictly covalent or contain a degree of ionicity, as is the case between Ga(-) ans As(+) in GaAs. Still, the atoms are constantly in motion, vibrating around their equilibrium lattice position, each atom a tiny harmonic oscillator. As the atomic vibrations are closely coupled through their common bonds, the atomic vibrations can be seen as part of larger lattice vibrations, which exist in several modes (see figure (3-1).

Similarly as with an electromagnetic field, each vibration mode can be quantized. A quantum of excitation in one mode is called a phonon, and each phonon can be characterized by a wave vector $\mathbf{q}$ and angular frequency ω. Like an electron or photon, an unconfined phonon can then be described by a (non-normalized) plane wave function $e^{i\mathbf{q}\cdot\mathbf{r}}$.

Similarly as for electrons, the lattice periodicity gives rise to a Brillouin-zone type $E - \mathbf{q}$ phonon dispersion relation. The lower branches represent the acoustic phonon
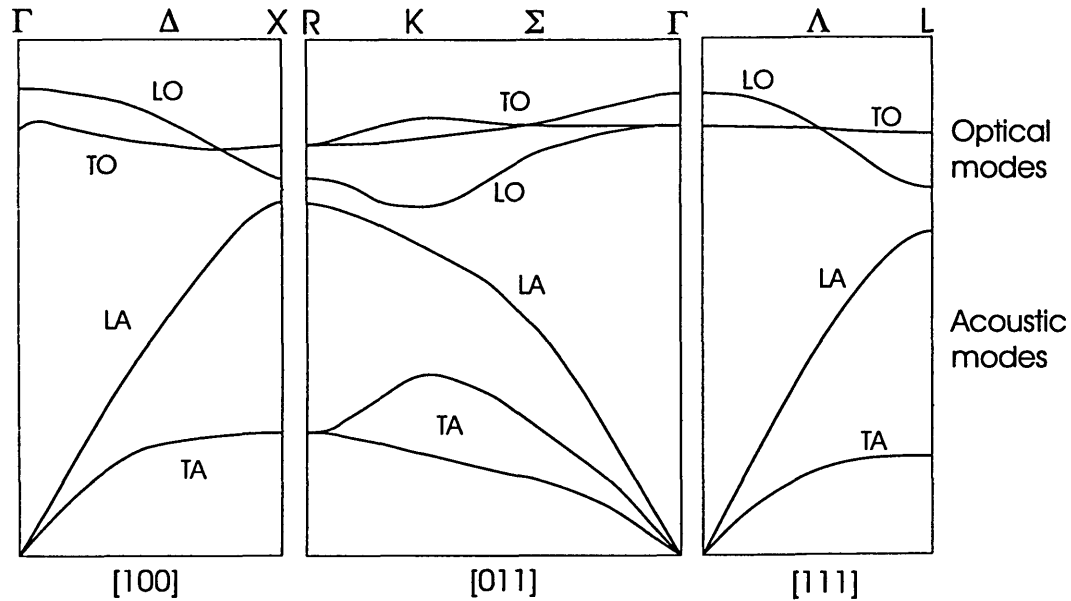
Figure 3-1: *Room temperature dispersion curves for acoustic and optical branch phonons in GaAs, obtained by inelastic neutron scattering. Adapted from Blakemore [25].*

modes, characterized by the neighboring atoms being in phase. In the longitudinal mode, the atomic displacements are in the same direction as the direction of energy transfer, while in the the transverse mode the atomic displacements are perpendicular to this direction. In optical phonon modes, the displacements of neighboring atoms are in opposite phase.

As shown in the figure, the energy of the optical phonons is almost independent of q, and in calculations longitudinal optical (LO) phonons are usually assumed to have one energy ($\hbar\omega_{LO} = 36$ meV in GaAs). Because of the large dipole moment induced by neighboring ions, LO phonons couple strongly with electrons in polar semiconductors, provided enough energy is available.

As the phonons themselves represent the motion of atoms which are centers of charge, they also represent time-dependent perturbations of the crystal potential and can therefore scatter charge carriers. The electron-phonon interaction, i.e. creation

and absorption of phonons, can be quantized through creation and absorption operators. The perturbation Hamiltonian H' is very similar to (2.76):

$$H' = \sum_{\mathbf{q}} \alpha(\mathbf{q}) \left[e^{-i\mathbf{q}\cdot\mathbf{r}} \, \mathbf{b}_{\mathbf{q}}^{\dagger} + e^{i\mathbf{q}\cdot\mathbf{r}} \, \mathbf{b}_{\mathbf{q}} \right]. \tag{3.1}$$

To find the total scattering rate, we sum over all $\mathbf{q}$ in the above equation. $\alpha(\mathbf{q})$ is the electron-phonon coupling strength.

LA phonons are most important for low energies or low temperatures. In these cases they correspond to long wavelength deformations of the crystal lattice. The interaction strength can be expressed as [29]:

$$|\alpha(\mathbf{q})|^2 = \frac{\hbar\omega D^2}{2\rho V c_s^2}. \tag{3.2}$$

Here D refers to the deformation potential, V is the crystal volume, ρ the material density and c_s the (longitudinal) speed of sound in the material.

For polar semiconductors such as GaAs, interactions with LO phonons are most important. The LO phonon interaction strength is [30]:

$$|\alpha(\mathbf{q})|^2 = \frac{\hbar\omega_{LO} e^2 (\epsilon_s - \epsilon_\infty)}{2\epsilon_s \epsilon_\infty \epsilon_0 V q^2}, \tag{3.3}$$

where ϵ_s and ϵ_∞ denote the relative permittivity of GaAs at frequencies lower and higher than optical frequencies, respectively. Due to the $1/q^2$ dependence, interactions with LO phonons at the zone center are favored over transitions involving a large momentum transfer. Also, the LO phonon threshold energy (36 meV for GaAs) causes a sharp temperature dependence. This effect is especially important for far-infrared transitions, where the subband energy spacing is less than $\hbar\omega_{LO}$. Here, hot carriers can open up parasitic LO phonon channels, drastically altering average scattering times. This can be detrimental to the working of the device.

We can adapt the bulk phonon expression to the 2D case by splitting the real space dependence of the Hamiltonian in components along and perpendicular to the growth axis.

$$H' = e \sum_{\mathbf{q}} \left[\frac{\hbar\omega_{LO} e^2 (\epsilon_s - \epsilon_\infty)}{2\epsilon_s \epsilon_\infty \epsilon_0 V q^2} \right]^{1/2} \frac{e^{-i\mathbf{q}_t \cdot \mathbf{r}_t}}{\sqrt{A}} \frac{e^{-i q_z z}}{\sqrt{L}}. \tag{3.4}$$

To assess the transition rate between an initial state $\mathbf{k}_i$ on subband i and a final state $\mathbf{k}_f$ on subband f, we use Fermi's Golden Rule:

$$W_{\mathbf{k}_i \to \mathbf{k}_f} = \frac{2\pi}{\hbar} H'_{if} \, \delta(E_i(\mathbf{k}_i) - E_f(\mathbf{k}_f) - \hbar\omega(\mathbf{q})). \tag{3.5}$$

We can rewrite the matrix element H'_{if}:

$$H'_{if} = \alpha(q) A_{if}(q_z) \, \delta_{\mathbf{k}_i - \mathbf{k}_f, \mathbf{q}_t}, \tag{3.6}$$

where $A_{if} = \displaystyle\int_{-\infty}^{+\infty} \psi_f^\dagger(z) \, \psi_i(z) \, e^{iq_z z} dz$.

The form factor A_{if} contains the dependence on the electron wave functions. Summing over all possible $\mathbf{k}_f$, we find the total rate for a LO phonon mediated transition from the initial state $\mathbf{k}_i$ in subband i to a state in subband f :

$$W_{\mathbf{k}_i} = \frac{2\pi}{\hbar} \sum_{\mathbf{k}_f} \sum_{\mathbf{q}_t} \sum_{q_z} |\alpha(q)|^2 \, |A_{if}(\mathbf{q})|^2 \, \delta(E_i - E_f - \hbar\omega(\mathbf{q})). \tag{3.7}$$

We assume parabolic subbands and $\hbar\omega(\mathbf{q}) \approx \hbar\omega_{LO}$. The electrons in subband i are thermalized with electron temperature $T_{e,i}$, and their energy distribution can be described using a Fermi-Dirac distribution around a chemical potential (Fermi energy) E_F:

$$f(E) = \frac{1}{e^{\frac{E - E_F}{kT_{e,i}}} + 1}. \tag{3.8}$$

As phonons are bosons, their energy distribution is the Bose-Einstein function:

$$N_{LO}(E) = \frac{1}{e^{\frac{E - \mu_{LO}}{kT_{ph}}} - 1}. \tag{3.9}$$

Using the momentum conservation $\mathbf{q}_t = \mathbf{k}_i - \mathbf{k}_f$ from (3.6) the LO phonon scattering rate can be written:

$$W_{\mathbf{k}_i} = \frac{e^2(\epsilon_s - \epsilon_\infty)\omega_{LO} \, m^*}{4\pi^2 \epsilon_0 \epsilon_\infty \epsilon_s \hbar^2} \int d\phi \int_0^{+\infty} dq_z \, (1 - f_f(E_f)) \frac{|A_{if}(q_z)|^2}{q^2}$$
$$\times (N_{LO}(\mathbf{q}) + 1) \, H(\varepsilon_i + \Delta E_{if} - \hbar\omega_{LO}). \tag{3.10}$$

ε_i is the kinetic energy of the electron in the initial subband, $\varepsilon_i = \hbar^2 k_{i,t}^2 / 2m^*$. The energy to create a phonon can be obtained from both the electron kinetic energy and potential energy, i.e. the subband separation. This is reflected in the step function $H(\varepsilon_i + \Delta E_{if} - \hbar\omega_{LO})$.

Usually we can assume that the phonon temperature T_{ph} is close to the lattice temperature T_l. For experiments where the device is being tested in a cryogenic environment, mounted on a cold plate, this means that LO phonons are frozen out. The equilibrium LO phonon population is negligible and scattering is dominated by the phonon emission process.

However, heavy pumping resulting in hot carriers could result in a non-equilibrium phonon population. The LO-phonon temperature T_{ph} can be much higher than the lattice temperature. This can result in absorption of LO phonons:

$$
\begin{aligned}
W_{k_i,abs} &= \frac{e^2(\epsilon_s - \epsilon_\infty)\omega_{LO}\, m^*}{4\pi^2 \epsilon_0 \epsilon_\infty \epsilon_s \hbar^2} \int d\phi \int_0^{+\infty} dq_z \left(1 - f_f(E_f)\right) \frac{|A_{if}(q_z)|^2}{q^2} \\
&\quad \times \left(N_{LO}(\mathbf{q})\right) H(\varepsilon_i).
\end{aligned}
\tag{3.11}
$$

The intersubband scattering time for the whole subband can then be found by averaging the individual scattering times of all electrons on that subband:

$$
W_i = \frac{1}{n_i} \sum_{\mathbf{k}_i} 2 f_i(E_i) W_{k_i},
\tag{3.12}
$$

where the factor 2 accounts for spin degeneracy, and n_i is the subband's 2D population density.

3.2 Carrier-carrier scattering

With increasing population density, electrons are more and more likely to interact and scatter. Especially in cases where LO-phonon scattering is not possible or very limited, e-e scattering is the main scattering mechanism. In this section we will be using the Hartree approximation, in which we neglect the "exchange energy" caused

$$22 \rightarrow 11 \qquad 22 \rightarrow 21 \qquad 21 \rightarrow 11$$
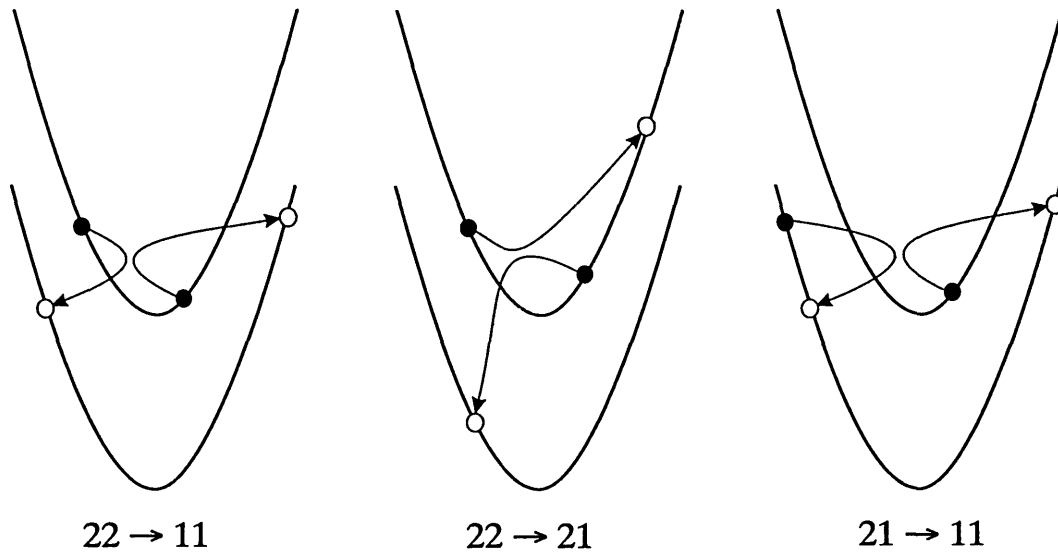
Figure 3-2: *Various intersubband carrier-carrier scattering mechanisms for a two-subband system*

by the anti-symmetry in the real space wave function of a two-electron state if the two electrons have the same spin. Inclusion of this "exchange energy" adds considerable complexity to the problem (Hartree-Fock), effectively making the problem intractable.

The perturbation Hamiltonian is an unscreened Coulombic potential:

$$H' = \frac{e^2}{4\pi\epsilon r}, \tag{3.13}$$

with r the distance between the electrons and $\epsilon = \epsilon_s\epsilon_0$ the dielectric permittivity of the semiconductor. At its simplest, we can represent electron-electron interaction as a two-body process involving two isolated carriers. As there are two initial and two final states, there are a lot more scattering possibilities than in the case of LO-phonon scattering, which involved one initial and one final state. In figure (3-2), various scattering mechanisms for intersubband scattering are illustrated. The transition from subband 2 to subband 1 can be split into three contributions, 22-11, 22-21 and 21-11. The 22-21 and 21-11 transitions are Auger-type transitions, with one electron relaxing down to a lower subband while giving its excess energy to another electron

$$22 \to 22 \qquad 21 \to 21 \qquad 21 \to 12$$
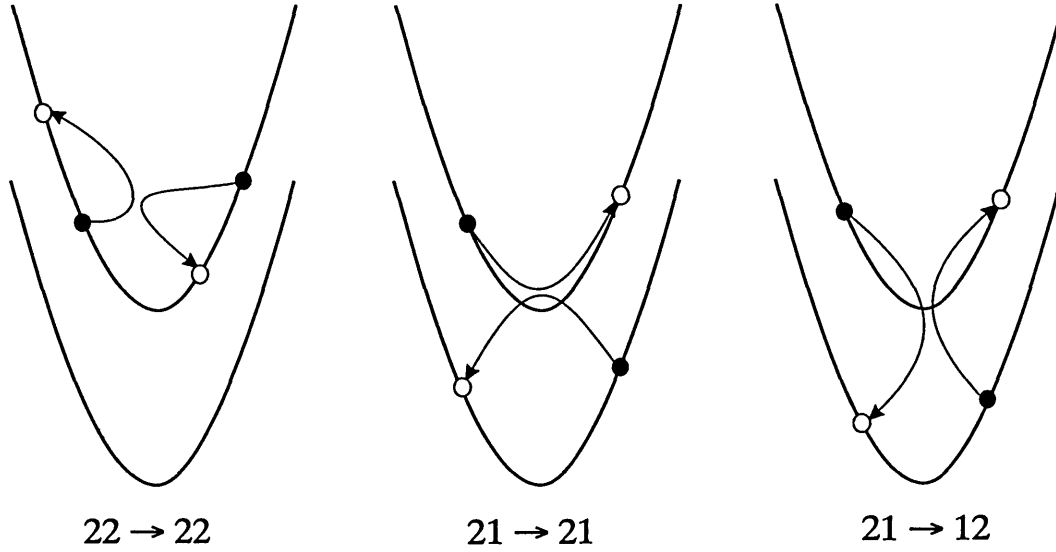
Figure 3-3: *Various intrasubband carrier-carrier scattering mechanisms for a two-subband system*

which scatters higher into its original subband. Also, there are scattering events which don't affect the number of electrons in a subband, as illustrated in figure (3-3). 22-22 is a "pure" intrasubband scattering event. Even though these intrasubband e-e scatterings don't change the subband populations, they are very important for thermal equilibrium in and between subbands.

The initial and final states are composed of two electron wavefunctions, and are of the form $|\psi_{12}\rangle = |\psi_1\rangle |\psi_2\rangle$. The collision probability between electrons with equal spin polarity is lower due to an exchange term (related to Pauli's exclusion principle), and therefore only electrons with opposite spins are taken into account here [31]. Taking wavefunctions of the form (2.93) and (2.94) for the electron wavefunctions, the matrix element H'_{if} becomes:

$$H'_{if} = \left\langle \psi_f(z)\frac{e^{-ik_f \cdot r_t}}{\sqrt{A}}\psi_g(z')\frac{e^{-ik_g \cdot r'_t}}{\sqrt{A}} \left| \frac{e^2}{4\pi\epsilon r} \right| \psi_i(z)\frac{e^{-ik_i \cdot r_t}}{\sqrt{A}}\psi_j(z')\frac{e^{-ik_j \cdot r'_t}}{\sqrt{A}} \right\rangle, \quad (3.14)$$

where the initial electron states are labeled i and j, and the final states f and g. Note that, for simplicity, we are working with the unscreened Coulombic potential.

The separation of the carriers is:

$$r = \sqrt{|\mathbf{r}_t - \mathbf{r'}_t|^2 + (z - z')^2}.$$ (3.15)

Therefore we obtain:

$$
\begin{aligned}
H'_{if} &= \frac{e^2}{4\pi\epsilon A^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int \int \psi_f^\dagger(z)\, \psi_g^\dagger(z')\, \psi_i(z)\, \psi_j(z') \\
&\times \frac{e^{i(\mathbf{k}_i \cdot \mathbf{r}_t + \mathbf{k}_j \cdot \mathbf{r'}_t)}\, e^{-i(\mathbf{k}_f \cdot \mathbf{r}_t + \mathbf{k}_g \cdot \mathbf{r'}_t)}}{\sqrt{|\mathbf{r}_t - \mathbf{r'}_t|^2 + (z - z')^2}}\, d\mathbf{r}_t\, d\mathbf{r'}_t\, dz\, dz'.
\end{aligned}
$$ (3.16)

Expanding the Coulombic potential in a Fourier series, and substituting H'_{if} into Fermi's Golden Rule gives the scattering rate of a carrier in subband i. Integrating over all the states of the second carrier (given by $\mathbf{k}_j$) and introducing Fermi-Dirac distribution functions to account for state occupancy, we find

$$
\begin{aligned}
W &= \frac{e^4}{2\pi\hbar(4\pi\epsilon)^2} \int \int \int \frac{|A_{ijfg}(q_t)|^2}{q_t^2} f_j(\mathbf{k}_j)\, [1 - f_f(\mathbf{k}_f)]\, [1 - f_g(\mathbf{k}_g)] \\
&\times \delta(\mathbf{k}_f + \mathbf{k}_g - \mathbf{k}_i - \mathbf{k}_j)\, \delta(E_f^t + E_g^t - E_i^t - E_j^t)\, d\mathbf{k}_g\, d\mathbf{k}_i\, d\mathbf{k}_j.
\end{aligned}
$$ (3.17)

where the energies E^t refer to the total energy of the corresponding carrier, i.e. subband minimum energy plus kinetic energy. A_{ijfg} is a form factor and a function of $q_t = |\mathbf{k}_i - \mathbf{k}_f|$:

$$A_{ijfg}(q_t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \psi_f^\dagger(z)\, \psi_g^\dagger(z')\, \psi_i(z)\, \psi_j(z') e^{-iq_t|z-z'|}\, dz'\, dz.$$ (3.18)

The delta functions express the conservation of momentum and energy in the scattering event. We can see that carrier-carrier scattering will be largest for small exchanged wavevectors. Assuming parabolic subbands with $E^t = E + \hbar^2 k^2 / 2m^*$, we find:

$$
\begin{aligned}
W &= \frac{m^* e^4}{\pi\hbar^3(4\pi\epsilon)^2} \int \int \frac{|A_{ijfg}(q_t)|^2}{q_t^2}\, P_{j,f,g}(\mathbf{k}_j, \mathbf{k}_f, \mathbf{k}_g) \\
&\times \delta\left(k_f^2 + k_g^2 - k_i^2 - k_j^2 + \frac{2m^*}{\hbar^2}(E_f + E_g - E_i - E_j)\right) d\mathbf{k}_f\, d\mathbf{k}_j,
\end{aligned}
$$ (3.19)
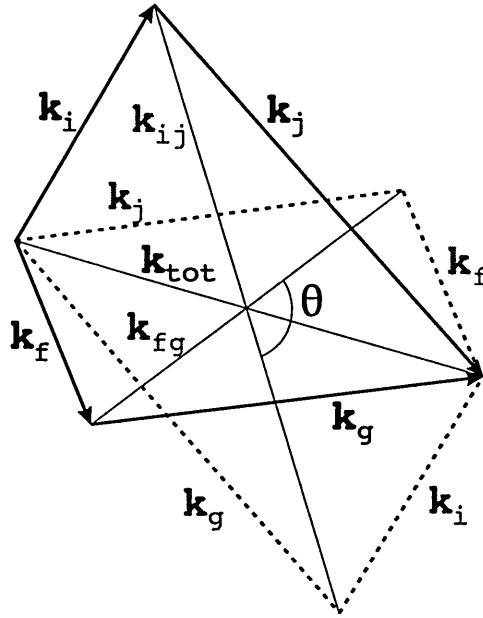
Figure 3-4: *Conservation of momentum in e-e scattering.*

with $P_{j,f,g}(\mathbf{k}_j, \mathbf{k}_f, \mathbf{k}_g)$ representing the probability functions. This equation is no longer valid for the non-parabolic valence band subbands in a quantum well.

It is useful to introduce two new variables, the relative wavevectors

$$\mathbf{k}_{ij} = \mathbf{k}_j - \mathbf{k}_i, \tag{3.20}$$

$$\mathbf{k}_{fg} = \mathbf{k}_g - \mathbf{k}_f. \tag{3.21}$$

The energy conserving delta function then allows reduction of this integral to:

$$W = \frac{m^* e^4}{\pi \hbar^3 (4\pi\epsilon)^2} \int \int_0^{2\pi} \frac{|A_{ijfg}(q_t)|^2}{q_t^2} P_{j,f,g}(\mathbf{k}_j, \mathbf{k}_f, \mathbf{k}_g) d\theta \, d\mathbf{k}_j, \tag{3.22}$$

and θ is the angle between $\mathbf{k}_{ij}$ and $\mathbf{k}_{fg}$, as illustrated in figure (3-4). Although (3.22) looks simple, the actual computation is rather time-consuming and resource-intensive. Certain simplifications, like ignoring final-state blocking, are common. This is a fair approximation for low carrier densities or high electron temperatures.

At higher electron densities, the interaction can no longer be described as solely between two isolated carriers. The reaction of other carriers to the Coulombic potential will effectively "screen" the disturbing field, thereby reducing the perturbation. The probability of scattering will decrease as compared to the non-screened case.

One of the simplest models for screening considers only the carriers within the same subband as the initial carrier state. It replaces [32] the dielectric constant ϵ_s with one which is dependent upon the relative wave vector q_t:

$$\epsilon_s = 1 + \frac{2\pi e^2}{(4\pi\epsilon_s)q_t}\Pi_{ii}(q_t, T)A_{ijfg}(q_t),$$ (3.23)

with the polarization factor

$$\Pi_{ii}(q_t, T) = \int_0^{+\infty} \frac{\frac{m^*}{\pi\hbar^2}\left[1 - H(q_t - 2k_F)\sqrt{1 - \left(\frac{2k_F}{q_t}\right)^2}\right]}{4kT\cosh^2(\frac{E_F - E}{2kT})}\, dE.$$ (3.24)

k_F is the Fermi wave vector for subband i.

Equation (3.22) gives the carrier-carrier scattering rate for a particular carrier energy i, averaged over another initial carrier distribution j. In order to find a scattering rate for the whole subband i, we have to average out over the Fermi-Dirac distribution of carriers in the initial state.

$$\frac{1}{\tau} = \frac{\int \frac{1}{\tau_i} f_i(\varepsilon_i)\, d\varepsilon_i}{\int f_i(\varepsilon_i)\, d\varepsilon_i},$$ (3.25)

with ε_i the kinetic energy associated with k_i. The denominator is equal to $N_i\pi\hbar^2/m^*$, and assuming a parabolic subband, we obtain:

$$\frac{1}{\tau} = \frac{\int \frac{1}{\tau_i} f_i(k_i)k_i\, dk_i}{\pi N_i}.$$ (3.26)

In figure (3-5) the temperature dependence of various inter- and intrasubband electron-electron scattering rates is shown. The "intrasubband" carrier scattering rate increases with temperature, as state blocking becomes less important. The rise in the number of easily accessible final states results in a higher scattering rate.
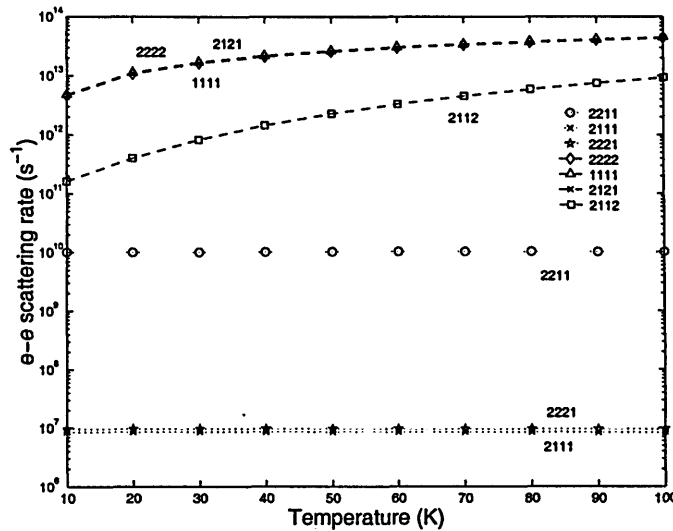
Figure 3-5: *Intrasubband and intersubband electron-electron scattering rates between the two lowest subbands of a 100 Å $Al_{.3}Ga_{.7}As/GaAs$ quantum well. Both subbands have a population density of $10^{10}cm^{-2}$. "Intersubband" processes are indicated with a dotted line, "intrasubband" processes with a dashed line.*

For intersubband scattering mechanisms this is less evident. Overall, intersubband e-e scattering can be considered to be nearly temperature insensitive. The 22-21 and 21-11 transitions are "forbidden" in this single quantum well, because their form factor vanishes [30]. The rate indicated in the figure is a numerical error. On the other hand, intersubband scattering is approximately proportional to the upper subband population at higher temperatures (figure (3-6)) when ignoring final state blocking. Inclusion of a significant population density in the final subband could lead to final state blocking being important, and hence give rise to a decrease in the scattering rate.

Figure (3-7) shows the dependence of e-e intersubband scattering on the intersubband separation. The different energy separations were obtained by varying the well width. The scattering time is very nearly inversely proportional to the intersubband energy separation. However, the relation becomes more complex if the two levels considered are close to anti-crossing. The overlap between the wavefunctions of initial
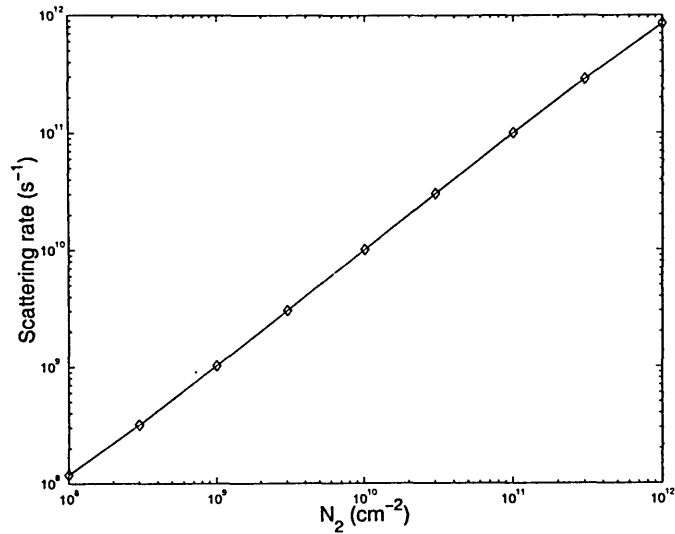
Figure 3-6: *Intersubband electron-electron scattering rates between the two lowest subbands of a 100 Å $Al_{.3}Ga_{.7}As/GaAs$ quantum well, as a function of the upper subband population.*
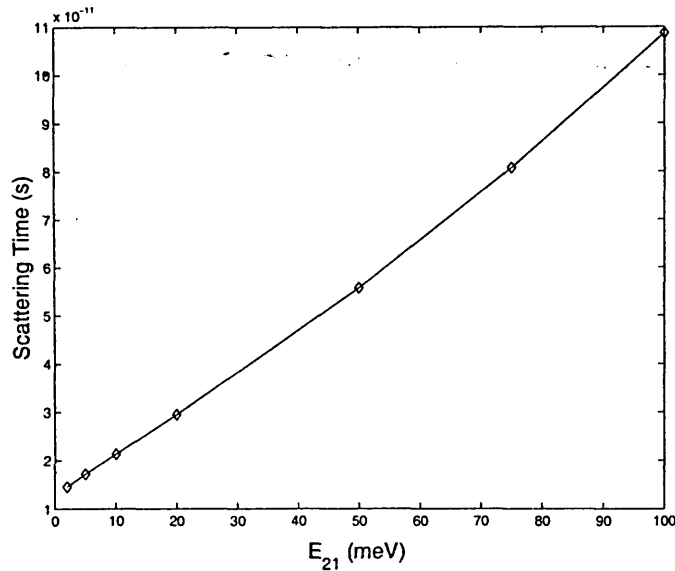
Figure 3-7: *Intersubband electron-electron scattering rates between the two lowest subbands in $Al_{.3}Ga_{.7}As/GaAs$ quantum wells of varying width, as a function of the subband energy separation. Both subbands have a population density of $10^{10}cm^{-2}$.*

and final states will sensitively depend on their energy difference.

We can summarize the above as:

$$\frac{1}{\tau} \propto \frac{N}{\Delta E_{21}}.$$

(3.27)

This empirical relation can be used as a rough guideline in the design of quantum well structures where carrier-carrier scattering plays an important role.

3.3 Electron Temperature in an Optically Pumped Device

The intense optical pumping in an interband pumped device dumps a lot of power into the device. Cooling the device down by mounting it on a cold finger reduces the heating, especially of the lattice (phonons). However, as electrons are generated several tens of meV above the band edge in some of the subbands, the excess kinetic energy of the electrons will cause the electron gas in each subband to heat up. The electron temperature can be considerably higher than the lattice temperature.

Carrier heating has been an active research topic for the past several years. Usually, two main subjects are addressed, thermalization and the energy loss rate of a thermalized carrier gas to the lattice. Researchers can use either band-to-band luminescence (most in undoped samples) or pump-and-probe techniques to measure the carrier energy distribution. Both pulsed and steady state experiments are important to understand the full picture. Once the electron or hole gas is thermalized, the high energy tail of the thermalized distribution can be used to deduce the carrier temperature.

The relaxation of carriers takes place through several different mechanisms. Assuming negligible carrier recombination, for low excess energy excitation the dominant thermalization mechanism is carrier-carrier scattering. Most carriers do not acquire enough energy to emit LO phonons. Experiments using high-energy excess have indicated relaxation processes where inter-valley scattering and sequential LO phonon emission play an important role.

Thermalization of the newly generated carriers by intrasubband scattering takes place within 300 fs to one ps for electrons [33] and less than 100 fs for holes [34]. Actual thermalization times depend on the carrier density and excess energy. For holes, the kinetic energy is never more than a couple of meV, and LO phonon scattering rarely comes into play. For electrons however, the excitation energy can come close to or

exceed the LO phonon energy. This can lead to accelerated inter- and intrasubband scattering due to interactions with LO phonons.

(a)   (b)
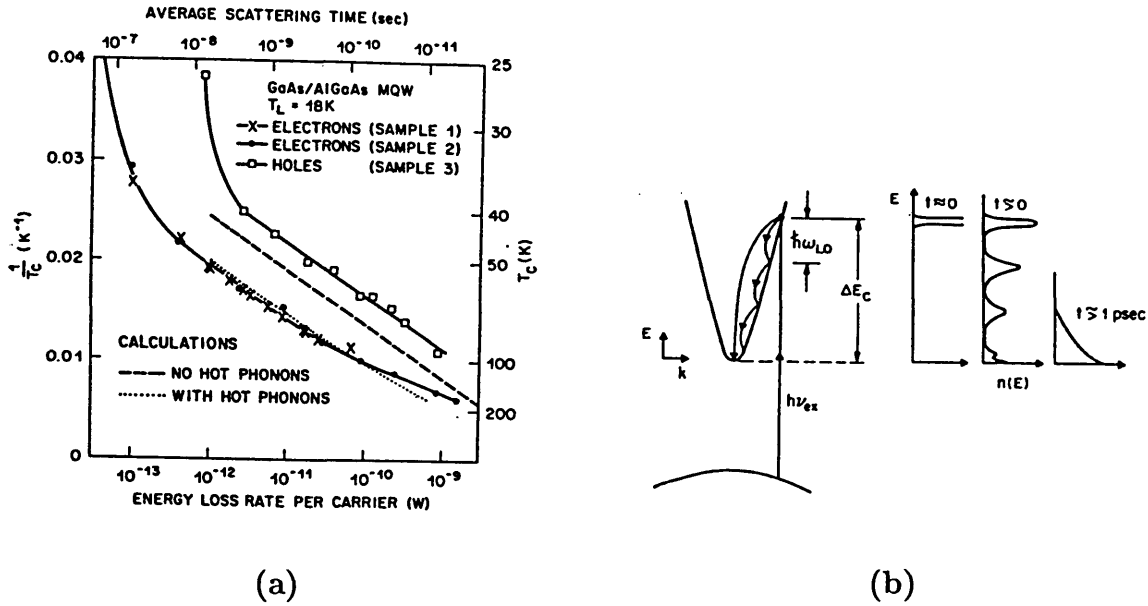
Figure 3-8: (a) Energy loss rates for holes and electrons for a MQW structure. At the same temperature, hole ELR's are one order of magnitude higher than electron ELR's [35]. (b) Schematic representation of the thermalization process after optical excitation [36].

The thermalized distribution then exchanges energy with the lattice (LO and LA phonons) and continues to cool down. The involved time constants range from tens (for holes) to hundreds of picoseconds (for electrons). In figure (3-8) we can see that the energy loss rates for electrons are about one order of magnitude smaller than for holes.

A typical experimental dependence of the electron temperature T_e on the optical pump intensity is shown in figure (3-9). Generally, the electron temperature in a subband will be in the range 20-100 K for most pump intensities. In the proposed quantum well structure, electron temperatures will probably be different in each conduction subband. This is due to the variations in the excess energy of the
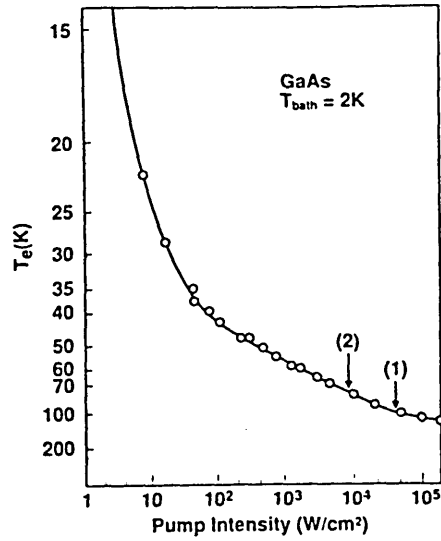
Figure 3-9: T_e versus excitation intensity for bulk GaAs at $E_{pump}=2.41$ eV, $T_{lattice}=2K$ [37].

photogenerated electrons, as well as the unidirectional flow of electrons from the upper laser level to the collection (depopulation) level. In the simulations, this difference is not taken into account. Instead, the temperature of all conduction and valence band levels is assumed to be equal. This assumption is justified because intrasubband e-e scattering, including "intersubband" e-e scattering with mere energy exchange, occurs on a time scale of 100 fs, as indicated in figure (3-5). Energy exchange between subbands is faster than intersubband particle exchange, which leads to an almost uniform temperature for all involved subbands. The calculations are then carried out for a set of temperatures between 20-100 K.

Chapter 4

Design and Simulation

4.1 Three-Level System

The simplest design is the "three-level" structure, illustrated in figure (4-1). Electrons are optically pumped from v_3 into the upper conduction band level (c_3). In order to ensure an efficient use of the pumping power, we will assume surface pumping. This means that the incident pump beam propagates along the QW growth axis, and the electric field is polarized in the plane of the well (TE polarization). The electrons can escape c_3 by either electron-electron (e-e) scattering down to c_2, LO-phonon scattering to c_1 or emission of a photon $\hbar\omega_{32}$. The transition $c_3 \rightarrow c_2$ is the target intersubband transition. The difference in energy between the lower intersubband laser level c_2 and c_1 is slightly greater than the LO phonon activation energy $\hbar\omega_{LO}$. This ensures a rapid LO-phonon-assisted depopulation of c_2, which is crucial in achieving a population inversion between the two far-infrared laser levels.

Unfortunately, in this design it is not possible to selectively pump c_3. In order to remove the optically generated holes from valence subband level v_3, the valence subband levels v_2 and v_1 need to have a lower or similar hole energy. This means that inevitably electrons will be optically excited from v_2 and v_1 into c_2 and c_1. Electrons and holes will gather in c_1 and v_1, respectively. If the carrier density in c_1 and v_1
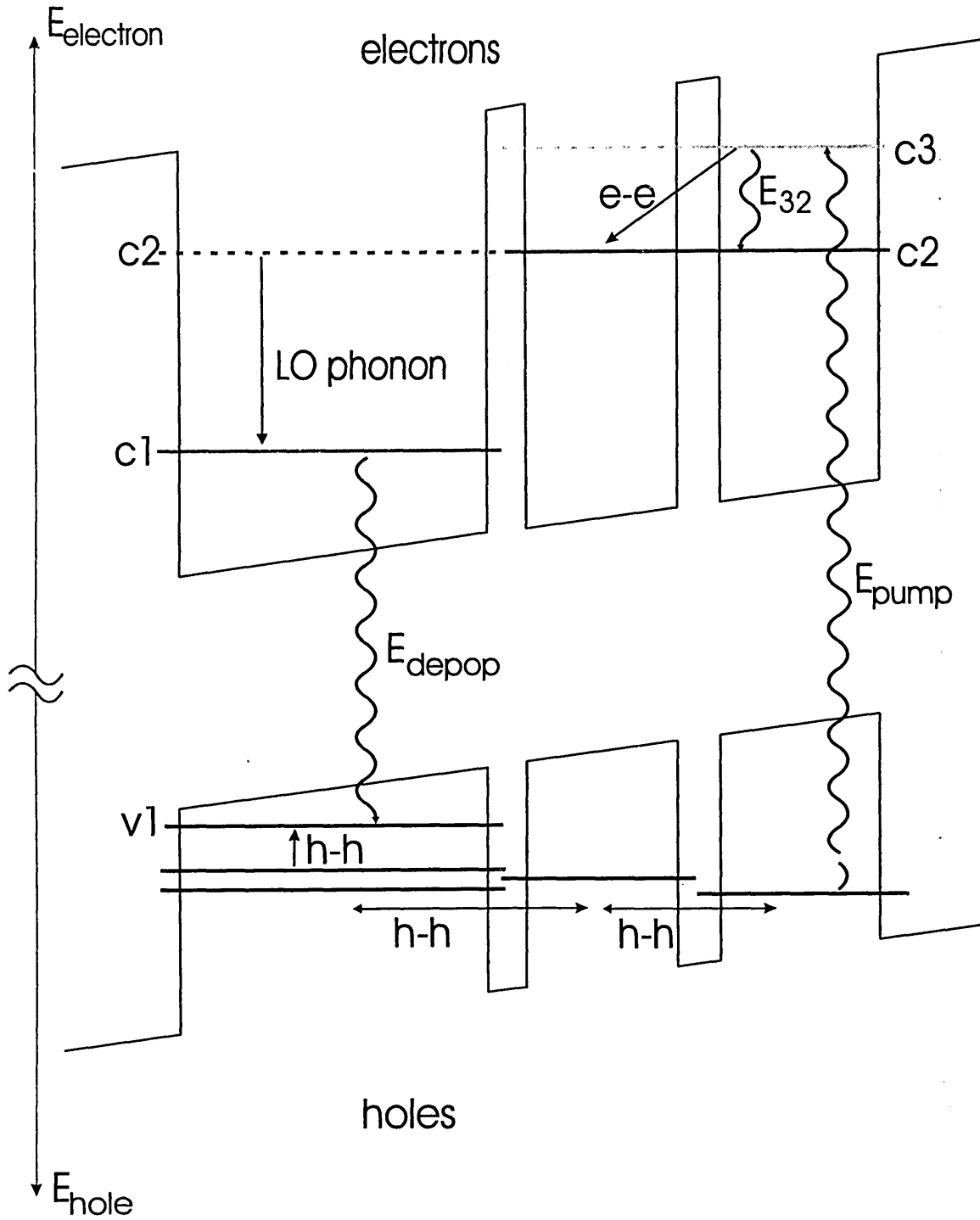
Figure 4-1: *Energy levels and transitions in a three-level system.*

becomes high enough, the lasing threshold for the interband transition $c_1 \rightarrow v_1$ is reached. Depopulation of the bottom levels of the valence and conduction bands by stimulated emission will then keep the population densities of both levels at a fixed value. The "depopulation laser" is important for the prevention of a buildup of carriers in the bottom levels, which causes state blocking of the depopulation of c_2 and an increased free-carrier absorption at THz frequencies. The interband laser is also the main "circulation pump", returning the optically excited electrons in the conduction band to the valence band.

In the valence band the heavy hole subbands are most important in hole transport. The HH ground states are the lowest in energy and will contain most holes after thermalization of the photogenerated carriers. In practice the HH ground states in the narrow wells are nearly lined up with the first excited HH state in the wide well. The fast scattering between these lined up subbands combined with an efficient h-h scattering from the excited state into the ground state in the wide well, ensures that most holes are concentrated in v_1. In the wide well, the LH ground state is very close to the first excited HH state, and can play a role in hole transport [38]. This means that the LH bands will play a small role in the interwell electron transport, as interwell scattering involving these higher energy states is considerably slower than scattering between the ground valence subbands, except in the wide well. However, their contribution in photogeneration is taken into account for every conduction subband.

The steady state subband populations are calculated using a rate equation analysis. Below the threshold for the interband depopulation laser, the rate equations for this "three-level" system are given by:

$$\frac{dn_{c3}}{dt} = G_{c3} - \left(\frac{1}{\tau_{c3c1}} + \frac{1}{\tau_{c3c2}} \right) n_{c3}, \tag{4.1}$$

$$\frac{dn_{c2}}{dt} = G_{c2} + \frac{1}{\tau_{c3c2}} n_{c3} - \frac{1}{\tau_{c2c1}} n_{c2}, \tag{4.2}$$

$$\frac{dn_{c1}}{dt} = G_{c1} - R_{c1} + \frac{1}{\tau_{c3c1}} n_{c3} + \frac{1}{\tau_{c2c1}} n_{c2}. \tag{4.3}$$

For the valence band, similar equations apply. Above the interband lasing threshold, n_{c1} and n_{v1} are clamped at the interband laser threshold population density. G_i and R_i refer to generation and recombination rates per unit area. The subband generation rate G_i is a function of the carrier density, carrier temperature, and the density of states in the involved subbands. Its dependence on population density and temperature is through the Fermi level energy (see equation (2.99)), which makes this set of equations highly nonlinear. The generation terms G_i include the contributions of all subbands participating in the photogeneration into subband i. For example, G_{c2} represents the transitions from all valence subbands (HH and LH) in all three wells into subband c_2. We assume very fast (less than 1 ps) intrawell intersubband scattering in the valence band, so that carriers generated in the excited and LH subbands relax into the HH ground state before participating in interwell scattering. Except in the lowest energy subband, we can neglect the interband (spontaneous) recombination terms R, as LO phonon and e-e scattering are much faster than radiative interband recombination (~ 1 ns).

In steady state all carrier densities are constant, so $\frac{dn_i}{dt} = 0$. Recombination equals generation in either band:

$$R_{c1} = R_{v1} = G_{c1} + G_{c2} + G_{c3} = G_{v1} + G_{v2} + G_{v3}, \tag{4.4}$$

where $1/R_{c1} = \tau_{c1v1}^{sp} \approx 1$ ns below the interband lasing threshold and decreases rapidly above the the threshold to keep n_{c1} constant.

Population inversion in steady state is found to be:

$$n_3 - n_2 = G_{c3}\, \tau_{c3}^{tot}\left(1 - \frac{\tau_{c2c1}}{\tau_{c3c2}}\right) - G_{c2}\, \tau_{c2c1}, \tag{4.5}$$

with

$$\frac{1}{\tau_{c3}^{tot}} = \frac{1}{\tau_{c3c2}} + \frac{1}{\tau_{c3c1}}. \tag{4.6}$$

This equation is easy to interpret if we keep in mind that in a system with a generation rate G and a carrier lifetime τ, the steady state population density is

$n = G\tau$. In order to obtain a population inversion, it is not sufficient that c_2 is depopulated faster than c_3. The depopulation of c_2 has to be fast enough also to get rid of the carriers optically excited into c_2 itself. c_1 and the valence band levels influence the generation terms. Electron and hole statistics and transport in these levels will determine carrier generation and injection into c_3 and c_2.

The total spontaneous emission power for the intersubband transition $c_3 \rightarrow c_2$ is given by

$$I_{em} = \frac{\hbar\,\omega_{32}\,n_{c3}}{\tau_{32}^{spon}} = \hbar\,\omega_{32}\frac{G_{c3}\,\tau_{c3}^{tot}}{\tau_{32}^{spon}}.$$

(4.7)

τ_{c3}^{tot} represents two major intersubband scattering processes for c_3, as seen in equation (4.6). Scattering into c_2 is dominated by e-e scattering, as the energy separation between the two levels does not allow for LO phonon scattering. But for the $c_3 \rightarrow c_1$ transition, LO phonon emission is allowed. This LO phonon channel can make it very difficult to achieve a substantial population inversion. Therefore, the suppression of the $c_3 - c_1$ LO phonon transition is a major design challenge. Matters are complicated further by the demand that the $c_3 \rightarrow c_2$ transition must have a sizable dipole moment, in order to have an efficient far-infrared emission. A strong overlap between c_3 and c_2, combined with an overlap between c_2 and c_1 for efficient LO phonon depopulation of c_2, will result in a fast LO phonon depopulation of c_3. In the three-level scheme, a compromise has to be found between the strength of the intersubband laser transition and the efficiency of the lower intersubband laser level depopulation.

For a single module, we can find the THz spontaneous emission efficiency by taking the ratio of the spontaneous emission and the absorbed incident pump power I_p in one module. The generation rate G_{ci} can be expressed as the product of an interaction cross-section σ_{ci} and the pump photon flux $n_{ph}c/n$, or:

$$G_{ci} = \sigma_{ci}n_{ph}\frac{c}{n} = \sigma_{ci}\frac{I_p}{\hbar\omega_p}\frac{c}{n},$$

(4.8)

where $\hbar\omega_p$ is the pump photon energy. This enables us to find for the spontaneous emission efficiency:

$$\eta_3 = \frac{I_{em}}{I_p} = \frac{\sigma_{c3}}{\sigma_{c1} + \sigma_{c2} + \sigma_{c3}} \frac{\hbar\omega_{32}}{\hbar\omega_p} \frac{\tau_{c3}^{tot}}{\tau_{32}^{spon}}. \tag{4.9}$$

This efficiency is the ratio of emitted THz power over the absorbed pump power per module. It is important to optimize the absorption efficiency by using many modules, so that the incident power would be nearly completely absorbed, and the total power output/input efficiency is close to the prediction of equation (4.9).

The emission efficiency depends critically on the ratio of the fastest non-radiative (LO) scattering time to the spontaneous emission lifetime. It is therefore very important to minimize the LO phonon scattering channels depopulating the upper laser level. Another important parameter is the pump selectivity $\sigma_{c3}/(\sigma_{c1} + \sigma_{c2} + \sigma_{c3})$, which is close to 1/3 in the three-level design. The calculated emission efficiency for the designed three-level structure is shown in figure (4-12), and is close to 10^{-7}. The dependence on pump intensity and temperature is discussed later.

4.2   Four-Level System

The main drawback of the simple three-level design is that the required strong coupling between the two radiative subband levels inevitably leads to a fast LO-phonon mediated depopulation of the upper level. In a four-level design we attempt to make the lower laser level considerably more susceptible to LO-phonon depopulation by increasing the overlap with level 1. By bringing the second energy level in the first well close to anticrossing with the lower laser level, this level is partly delocalized. Fast LO-phonon scattering depopulates the lower intersubband laser level. However, there are several disadvantages to this scheme. The intersubband dipole moment is now split between two heavily mixed bands (4-3 and 4-2), which means the available gain will be split between two transitions as well. This decreases the maximum

available gain, and increases the intersubband laser threshold population inversion. Also, we assume that the electrons in the two mixed bands can tunnel coherently into the wide well, even though the barrier between the wide well and the adjacent well is fairly thick and could cause loss of coherence. In that case the electrons get trapped behind the barrier and have to tunnel through in order to scatter down, greatly reducing scattering times and increasing the population density in the lower FIR laser levels. However, in this non-coherent case the loss in oscillator strength is also reduced because of the effective localization of the electron wavefunctions in the two narrow wells.

Similarly as for the three-level case, we can write the steady state rate equations:

$$\frac{dn_{c4}}{dt} = G_{c4} - \left(\frac{1}{\tau_{c4c1}} + \frac{1}{\tau_{c4c2}} + \frac{1}{\tau_{c4c3}}\right) n_{c4}, \tag{4.10}$$

$$\frac{dn_{c3}}{dt} = G_{c3} + \frac{1}{\tau_{c4c3}}n_{c4} + \frac{1}{\tau_{c2c3}}n_{c2} - \left(\frac{1}{\tau_{c3c1}} + \frac{1}{\tau_{c3c2}}\right) n_{c3}, \tag{4.11}$$

$$\frac{dn_{c2}}{dt} = G_{c2} + \frac{1}{\tau_{c4c2}}n_{c4} + \frac{1}{\tau_{c3c2}}n_{c3} - \left(\frac{1}{\tau_{c2c1}} + \frac{1}{\tau_{c2c3}}\right) n_{c2}, \tag{4.12}$$

$$\frac{dn_{c1}}{dt} = G_{c1} - R_{c1} + \frac{1}{\tau_{c4c1}}n_{c4} + \frac{1}{\tau_{c3c1}}n_{c3} + \frac{1}{\tau_{c2c1}}n_{c2}, \tag{4.13}$$

and

$$R_{c1} = G_{c1} + G_{c2} + G_{c3} + G_{c4}. \tag{4.14}$$

Again, $1/R_{c1} = \tau^{sp}_{c1v1} \approx 1$ ns below the interband lasing threshold, and decreases rapidly above the threshold to keep n_{c1} at its threshold value. Here we neglected the recombination terms in all except the lowest subbands, as in the three-level case.

In figure (4-3) the dependence of the interband spontaneous lifetime on subband population density and electron temperature is illustrated. The lifetime was determined by taking the ratio of the spontaneous recombination rate and the population density. The quasi-linear dependence of the spontaneous lifetime on temperature can easily be seen if we assume a Boltzmann distribution of the carriers in the subband,
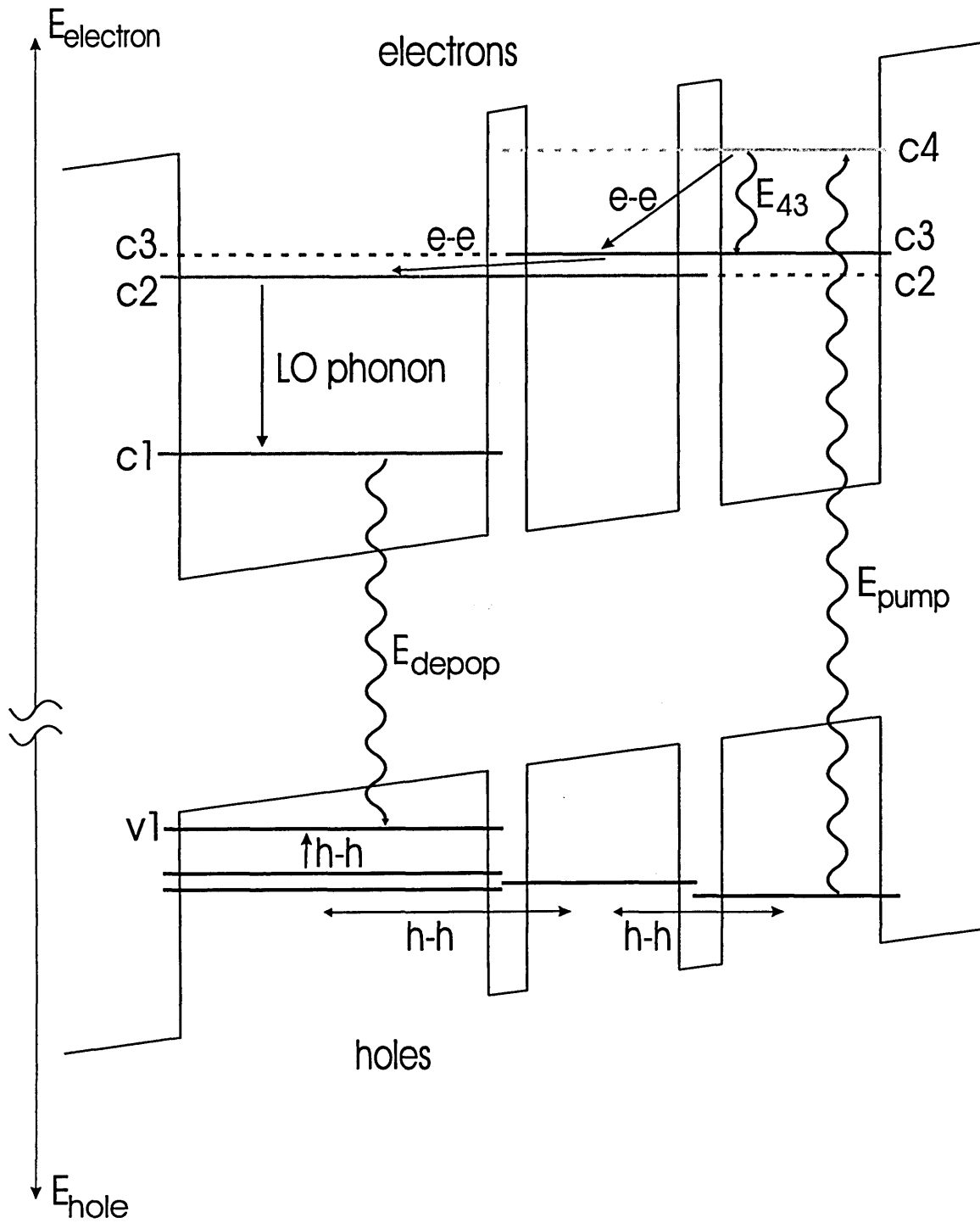
Figure 4-2: *Energy levels and transitions in a four-level system. The first excited subband in the wide well and the lower laser level are heavily coupled.*

(a)   (b)
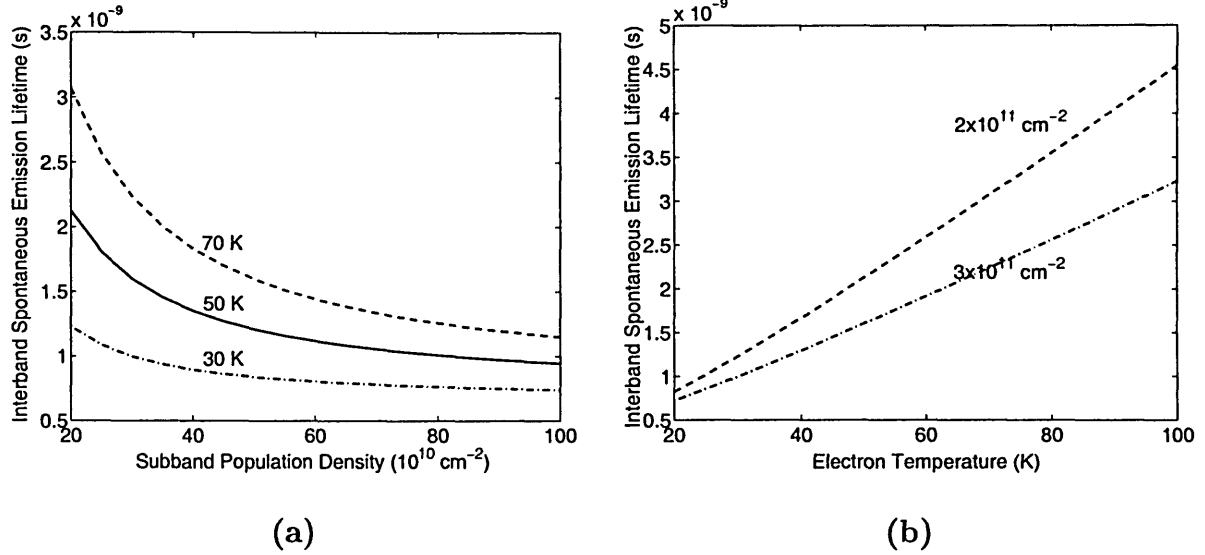
Figure 4-3: (a) *Dependence of the spontaneous interband emission time* τ_{clvl}^{sp} *on the population density in the laser levels for various electron temperatures.* (b) *Dependence of the spontaneous interband emission time* τ_{clvl}^{sp} *on the electron temperature for a population density of* $2 \times 10^{11} cm^{-2}$ *and* $3 \times 10^{11} cm^{-2}$.

and a constant density of states in the valence and conduction bands. We can then write for the population density N

$$N = g^{2D} \int_{E_0}^{\infty} e^{\frac{E_f - E}{k_B T}} dE, \tag{4.15}$$

with g^{2D} the 2D density of states, E_0 the band edge energy, and E_f the subband quasi-Fermi-energy. From this we find that $\exp(E_f/(k_B T)) \sim N/(k_B T)$. The spontaneous emission rate at a certain emission energy is proportional to the joint density of states for that energy, and to the number of electrons (in the conduction band) and holes (in the valence band) at that energy. The density of states is assumed constant, so we can leave it out of what follows, and for the electron and hole distribution we use the Maxwell-Boltzmann approximation. The spontaneous emission rate R_{ib}^{sp}

$$R_{ib}^{sp} \propto \int_{0}^{\infty} e^{-\frac{E_c(k) - E_{f,c}}{k_B T}} e^{\frac{E_v(k) - E_{f,v}}{k_B T}} dE, \tag{4.16}$$

where the indices c and v refer to conduction and valence band respectively, and the integration is over the interband transition energy in excess of the bandgap energy. We find that

$$R_{ib}^{sp} \propto e^{-\frac{E_{f,c}-E_{f,v}}{k_B T}} \int e^{\frac{E_v(k)-E_c(k)}{k_B T}} dE \propto \frac{1}{k_B T}. \tag{4.17}$$

Radiative interband transitions conserve in-plane electron momentum. As temperature rises, electrons (in the conduction band) and holes (in the valence band) get more spread out and have fewer counterparts available to recombine with, hence increasing the radiative recombination lifetime.

The dependence of the spontaneous interband lifetime on population density is closely related to the above. At low densities or higher electron temperatures, the carrier distribution function is spread out and somewhat resembles a Boltzmann distribution. Higher temperatures or lower densities will yield a longer lifetime. For lower temperatures or higher densities, the Fermi-Dirac distribution increasingly resembles a step function. In this limiting case, the spontaneous emission lifetime is minimal as there is no state-blocking.

The steady state expression for population inversion between subbands c_3 and c_4 can easily be deduced from the above equations:

$$n_{c4}-n_{c3} = G_{c4}\,\tau_4^{tot} - \left[G_{c3} + G_{c4}\left(\frac{\tau_{c4}^{tot}}{\tau_{c4c3}} + \frac{\tau_{c4}^{tot}}{\tau_{c4c2}}\frac{\tau_{c2}^{tot}}{\tau_{c2c3}} \right) + G_{c2}\frac{\tau_{c2}^{tot}}{\tau_{c2c3}} \right] \frac{\tau_{c3}^{tot}}{1 - \frac{\tau_{c3}^{tot}}{\tau_{c3c2}}\frac{\tau_{c2}^{tot}}{\tau_{c2c3}}}.\tag{4.18}$$

The second term represents n_3. The factor in brackets is a generalized generation term, encompassing the carriers scattering into c_3 from c_4 (directly and via c_2) and from c_2. The second factor is an effective scattering time. The denominator is an enhancing factor for τ_{c3}^{tot}, accounting for a feedback loop when electrons scatter from c_3 to c_2 and back.

Similar to the three-level case, the emission efficiency of the four-level intersubband spontaneous transition 4-3 can be written as:

$$\eta_4 = \frac{\sigma_{c4}}{\sigma_{c1} + \sigma_{c2} + \sigma_{c3} + \sigma_{c4}} \frac{\hbar\omega_{43}}{\hbar\omega_p} \frac{\tau_{c4}^{tot}}{\tau_{43}^{spon}}. \tag{4.19}$$

Again, the output/input efficiency will be close to equation (4.19) if all the incident power is absorbed in the active region. The calculated emission efficiency for the designed four-level structure is shown in figure (4-12), and is close to 10^{-7}.

4.3 Simulation

4.3.1 Computational model

The rate equations governing the subband populations and, subsequently, intersubband gain are highly coupled and nonlinear because the e-e and e-LO-phonon scattering rates depend on the density and temperature of the electrons. There is no analytical solution to describe the subband populations as function of the optical pump intensity. However, a steady state simulation of the system can yield a self-consistent solution to the problem. This means that, starting from an initial guess for the subband populations, the band potential profiles and scattering times are calculated. Using those parameters, new subband populations are determined. However, now these new values do not correspond to the previously calculated carrier-density dependent parameters, so we iterate the process until the solution converges. The convergence limit is the self-consistent solution.

The model incorporates the major scattering and optical transition processes expected to occur in the structure. There are two main parts in the simulation: the calculation of the valence and conduction band structure and the semi-self-consistent calculation of the subband populations.

The primary tool used in the design of the quantum wells is a program written by Paul Harrison (University of Leeds) [39]. This numerical simulation program iteratively solves Schrödinger's equation and Poisson's equation for a series of heterostructures under bias. First Schrödinger's equation is solved to determine the bound states, and then they are populated according to the expected carrier density. Poisson's equation is then solved to determine the effect of the accumulated charge
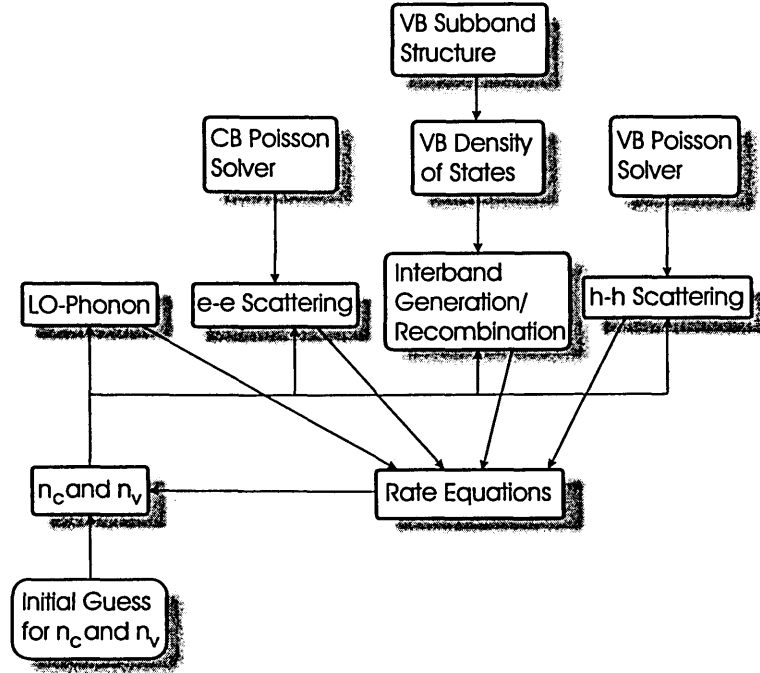
Figure 4-4: *Schematic overview of the different steps in the simulation of an interband pumped intersubband THz laser.*

on the shape of the potential wells. Schrödinger's equation is then re-solved with this new potential profile. This process is repeated until the result converges.

For each valence band well, the number of subbands and their $E - k$ dispersion relations are calculated. This is done using the two-band $k \cdot p$ formalism outlined in chapter 2.2-2.3. From the dispersion relations, we can deduce the density of states per subband. The next step is the calculation of the optical transition strength for the various interband transitions. This will determine the generation rate of photo-excited carriers in each subband and the spontaneous interband radiative lifetime $\tau_{c1v1}^{sp} = 1/R_{c1}$.

In the second part of the simulation, the subband population densities are calculated. The whole calculation is repeated for several carrier temperatures. For the first calculation, we use an initial guess for the different n_c and n_v. LO phonon scattering

rates are calculated for all intraband transitions where one can reasonably expect a LO phonon channel to open up and be important at the simulated temperature. Electron-electron scattering is usually only important where LO phonon scattering is not available. In the valence band, hole-hole scattering is the main means of hole transport. Carrier-carrier and phonon scattering calculations are performed with a program written by Paul Harrison [39]. The hole-hole scattering rates are approximated by treating them as if the valence subbands were parabolic bands with the appropriate carrier mass, neglecting the valence subband interaction. This crude approximation is justified because only an order-of-magnitude estimate of the hole-hole scattering rates is needed to roughly guess the hole distribution, and determine whether the great majority of the holes will rapidly relax into v_1.

For the calculation of the optical pumping rates, we assume thermalized subband populations. Using the previously calculated optical transition strengths, generation rates for the pump laser are found. Plugging the various generation rates and lifetimes into the rate equations, we find new values for the subband populations. Again assuming thermalized populations, the degree of population inversion between the lowest subbands 1 in conduction and valence band will decide whether the lasing threshold for the interband laser is reached (see figures (4-10) and (4-22)). If so, the population densities in these subbands are restricted to their threshold level. The new values for the subband populations can now be used to reiterate the above process, obtaining more consistent values with every iteration. Provided that the initial guesses for the subband populations are not too far off the calculated values, the impact on the solution of Schrödinger's equation will be minimal. Therefore it is sufficient to reiterate the rate equations.

Finally, the self-consistent population densities yield the expected gain for the intersubband THz laser.
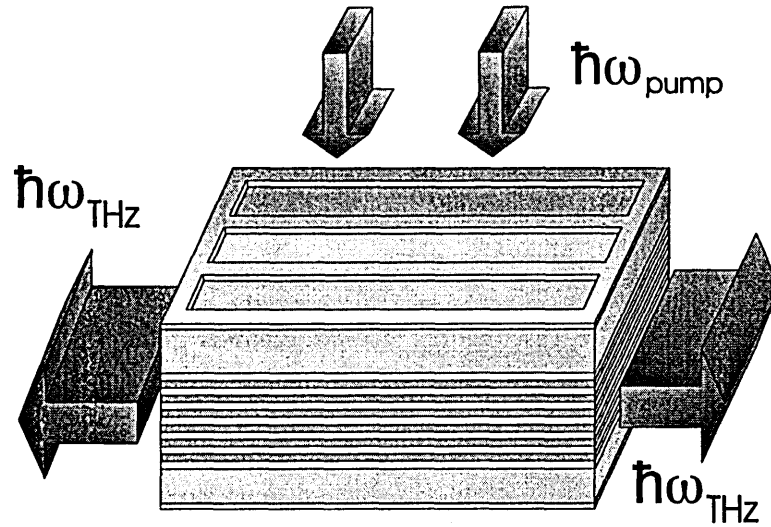
Figure 4-5: *Sketch of the surface pumped THz laser. A gold ridge window structure allows the pump beam to penetrate the sample while still providing a metal waveguide for the THz laser mode.*

4.3.2   Waveguide

In a laser, an electromagnetic wave propagates through a wave guide, containing both an active region (with net gain) and a cladding region, where losses prevail. The top and bottom of the waveguide are covered in gold to better confine the propagating mode. The top gold layer consists of a windowed structure with the length along the ridge. This orientation will not leak out the desired THz mode, but still allows for surface pumping. Details of the fabrication of such a waveguide are outlined in Bin Xu's thesis [40].

The active region consists of a number of quantum well modules, optimized to make an efficient use of the incident power of the pump beam and to maximize the confinement factor. The confinement factor denotes the overlap of the propagating mode with the active region, and is important to determine how much of the mode is actually useful in the process of stimulated emission. A schematic sketch of the surface pumped structure is shown in figure (4-5).

The cladding layers consists of a semiconductor material with a bandgap larger than the pump energy in order to avoid absorption in the cladding layer, and a dielectric constant lower than in the active region to help the optical mode confinement. The mirrors on both sides of the cavity are cleaved facets. The reflectance at one such mirror is:

$$R = \left(\frac{n_s - 1}{n_s + 1}\right)^2, \tag{4.20}$$

with n_s the refractive index of the semiconductor. For GaAs n_s=3.27 and we find $R = 0.28$.

4.3.3 Design Simulation

Design Parameters

One four-level module is depicted in figure (4-6). The energy levels are indicated in table (4.1). At 4K, E_g for GaAs is 1.52 eV. The THz laser levels are separated by 18.7 meV (4.5 THz or 67 μm) with a transition dipole moment of 36 Å. The energy differences between c_2 or c_3 and c_1 are higher than the LO phonon resonance, 36 meV. This is necessary to prevent state blocking of the LO-phonon scattering as the lowest conduction subband has a population density equal to the interband depopulation laser threshold density. This ensures fast LO phonon depopulation of the lower intersubband laser levels. The threshold density calculations are explained in more detail below.

A bias voltage of 10 mV/module helps to align the levels in both conduction and valence band. The bias is chosen to line up the valence subbands in a configuration that will enhance h-h scattering and efficiently transport the photogenerated holes into v_1. Between each module is a 28 ML barrier, preventing interaction between modules. Including this barrier, a module spans 139 monolayers, or 38.7 nm.

For comparison, a three-level module is shown in figure (4-7). The laser transition is 17 meV (4.1 THz or 73.2 μm) with a dipole moment of 36 Å, and a bias voltage of

18 mV/module is applied. This module is 121 ML (34.2 nm) wide.

Note that the radiative dipole moment for the three- and four-level system are very similar, despite the split-up of the four-level dipole moment due to the doublet lower subband. In the four-level design, the symmetry properties of the involved wavefunctions result in an enhancement of z_{43} over z_{42}. The dipole moment contributions originating from the wide well have an opposite sign for the 4-2 and 4-3 transitions, whereas the parts due to the overlap of the wavefunctions in the narrow wells are about equal for both transitions.

The 4-3 transition features both a larger dipole moment and a slower LO phonon depopulation than the 4-2 transition (figure (4-13)). This clearly qualifies the 4-3 transition as the better candidate for THz lasing. The slower LO phonon depopulation is mainly due to the smaller energy separation between c_4 and c_3, which means fewer electrons in c_4 will have sufficient energy to emit an LO phonon and scatter down to c_3 than to c_2.

In order to find the total thickness of the active region, we have to take the pump beam depletion into account. Calculation reveals that each conduction band energy level, available for interband absorption at an energy less than the pump energy, typically absorbs about 0.5-1% of the incident power.

The calculated absorption spectra for the conduction subbands in both the three- and four-level system are shown in figure (4-8). The pump energy is chosen so as to maximize photogeneration in c_4, but it is kept as close to the subband edge as possible to keep the electron temperature down. In the four-level case that is about 60 meV above the onset energy for the c_4 absorption, which is around 1600 meV (745 nm). The incident pump field is assumed to be multimodal with a Lorentzian line-shape envelope and a linewidth of 4 meV. There is no selectivity in the photogeneration at this pump wavelength, and the absorption coefficient for each conduction subband is around 2×10^3 cm^{-1} (0.78%/cond.subband). The total absorption coefficient is approximately 8×10^3 cm^{-1}, close to the bulk value of $\sim 10^4$ cm^{-1} [41]. For the three-
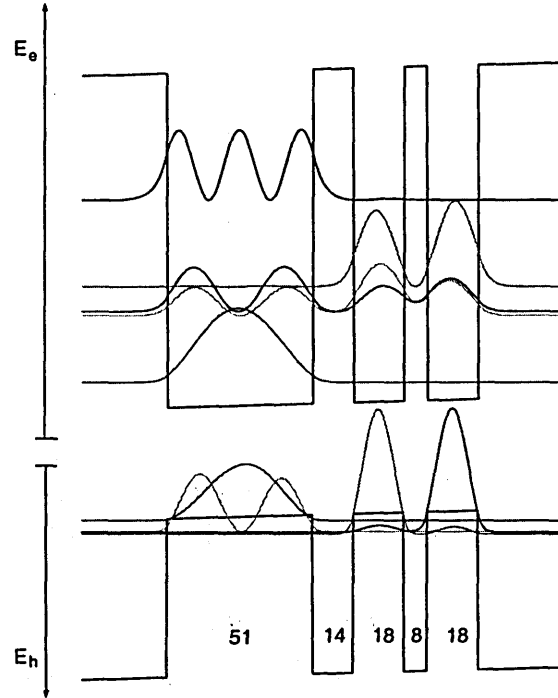
Figure 4-6: *Design parameters and energy levels for the four-level system. The well and barrier widths are indicated in units of monolayers (2.825 Å). A bias voltage of 10 mV/module is applied to help align the subbands.*

Subband Energy (meV)

E_{c1}	15.774	E_{v1}	$-E_g$-4.587
E_{c2}	66.010	E_{v2}	$-E_g$-13.356
E_{c3}	69.003	E_{v3}	$-E_g$-13.566
E_{c4}	87.671	E_{v4}	$-E_g$-15.289

Energy Separation (meV)

	E_{c1}	E_{c2}	E_{c3}
E_{c2}	50.24		
E_{c3}	53.23	2.99	
E_{c4}	71.90	21.66	18.67

Dipole Moment (Å)

	E_{c1}	E_{c2}	E_{c3}
E_{c2}	21.81		
E_{c3}	24.06	83.6	
E_{c4}	1.46	18.49	35.72

Table 4.1: *Energy levels for the module shown in figure (4-6).*
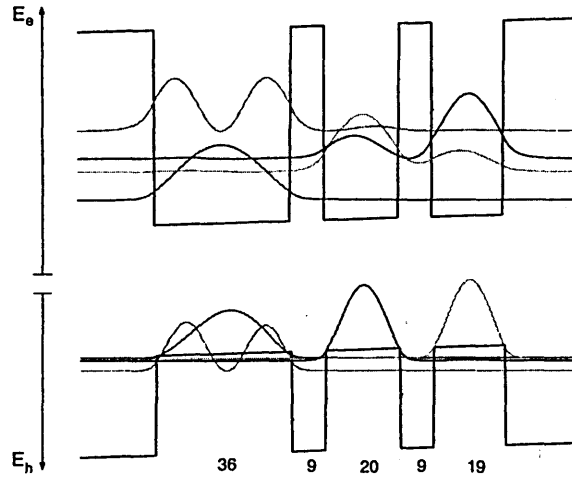
Figure 4-7: *Design parameters and energy levels for the three-level system. The well and barrier width are indicated in units of monolayers (2.825 Å). A bias voltage of 18 mV/module is applied to help align the subbands.*

Subband Energy (meV)			
E_{c1}	25.794	E_{v1}	$-E_g$-9.093
E_{c2}	62.657	E_{v2}	$-E_g$-9.850
E_{c3}	79.675	E_{v3}	$-E_g$-12.686

Energy Separation (meV)		
	E_{c1}	E_{c2}
E_{c2}	36.86	
E_{c3}	53.88	17.02

Dipole Moment (Å)		
	E_{c1}	E_{c2}
E_{c2}	11.888	
E_{c3}	6.211	36.452

Table 4.2: *Energy levels for the module shown in figure (4-7).*
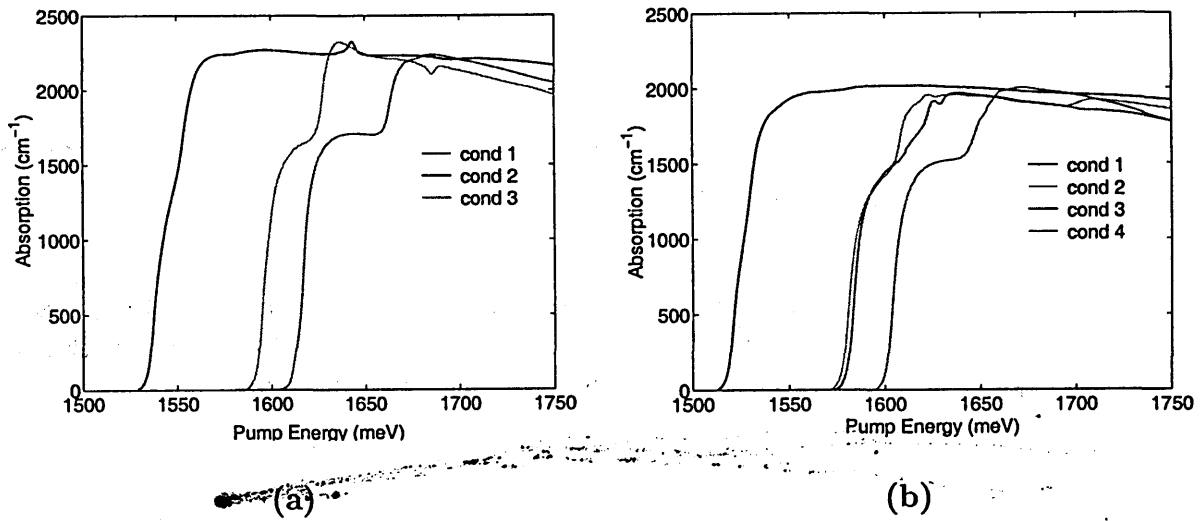
Figure 4-8: (a) *Absorption as a function of the incident photon energy for each of the considered subbands in the conduction band. Each absorption curve consists of contributions from the various valence subbands. (a) refers to the three-level system, (b) shows the absorption for the four-level case. The difference in absorption is due to the smaller module length in the three-level case, resulting in a higher 3D density of states.*
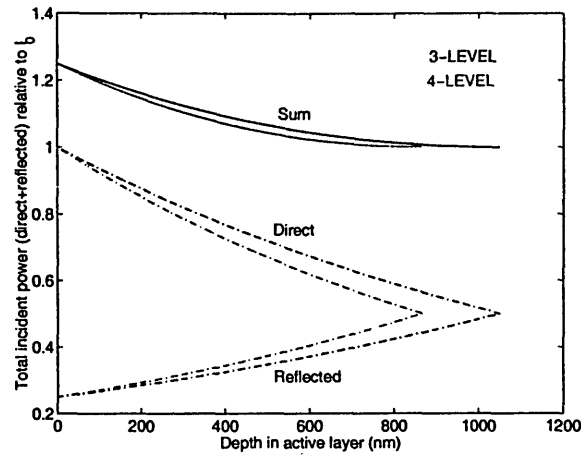
Figure 4-9: *Pump beam intensity in the active region for both 3- and 4-level designs. The beam is reflected at the point where its intensity is 50 % of its original value.*

level structure, the active layer thickness is 1050 nm, corresponding to 30 modules.

This allows us to make a rough estimate of the maximum number of modules. The incident power is absorbed in the modules and decays exponentially with a penetration depth of 1.25 μm. If we allow the pump beam to reflect off the back mirror when its intensity drops-to 50%, we find for the width of the active layer:

$$w = -\ln(0.5)/8000 \text{ cm} = 866 \text{ nm},$$

corresponding to 22 modules. For the three-level structure the absorption is close to 2.2×10^3 cm^{-1} per conduction subband. The active layer thickness is 1050 nm, corresponding to 30 modules. An approximate transverse intensity profile for the pump beam is shown in figure (4-9).

Typical device surface area dimensions are of the order of 500μm$\times$20μm. Simulations are done for a range of pump powers going from 50 mW to 50 W concentrated on the surface of the device. This corresponds to power densities of 500 W/cm^2 to 500 kW/cm^2.

For the calculation of the threshold population density of the interband "depopulation" laser transition, we take the mirror loss to be 23 cm^{-1}, equal to the mirror loss
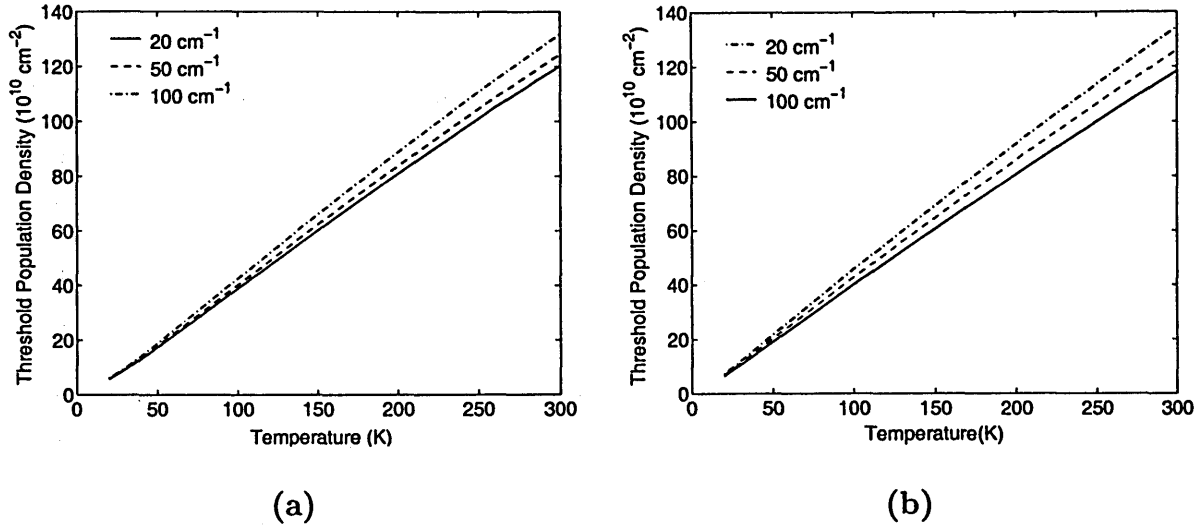
Figure 4-10: *Temperature dependence of the threshold population density for the interband laser in the* (a) *three- and* (b) *four-level systems. Indicated are threshold densities for a round-trip cavity loss of 20, 50 and 100 cm^{-1}.*

in a 500 μm cavity, $\alpha_m = \frac{1}{L} \ln R$. Taking into account various other loss mechanisms, mode scattering, mode leakage and absorption in the gold layers, the cavity loss is estimated to be around 50 cm^{-1}.

The population density of the valence subband and conduction subband is taken to be equal. Because these subbands are the lowest in energy in the module, the majority of the carriers gathers in these levels, scattering down from the other subbands. The carrier concentration in these wells is limited by the threshold carrier density for the interband laser. Assuming thermalized carrier populations, the threshold n_{c1} and n_{v1} can be found by using the gain equation (2.99). A plot of the threshold population density versus carrier temperature for different values for the threshold gain is shown in figure (4-10). We see that at a carrier temperature of 50 K, threshold densities in the lowest valence and conduction subbands are close to 2×10^{11}cm^{-2}, corresponding to a quasi-Fermi level of about 7 meV above the conduction subband edge. In order to avoid state blocking for the LO phonon scattering from the lower intersubband laser level to the c_1 subband, the intersubband spacing between these levels has to be
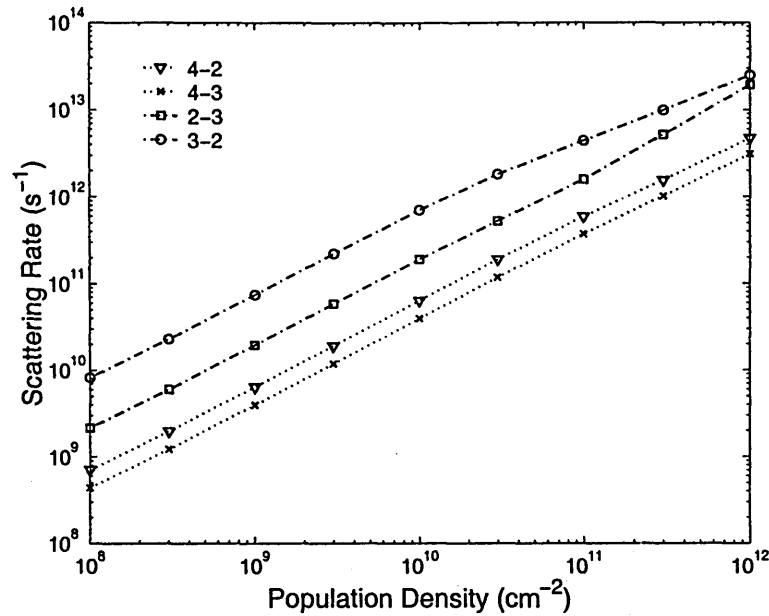
Figure 4-11: *Electron-electron scattering rates vs. carrier density for the four-level system. The scattering rates vary almost linearly with the population density. The electron temperature is 50K.*

over 43 meV.

Electron-electron scattering rates in the four-level system are shown in figure (4-11). Again we see that the e-e scattering rates vary nearly linearly with the subband population density. Remarkable is the large difference between the 3-2 and 2-3 scattering rates, though both transitions have very similar form factors. However, the fast 22-33 and 22-23 scattering processes are partially suppressed because of energy conservation for electrons close to the c_2 band edge. In the expected range of population densities for c_2 and c_3, i.e. $10^8 - 10^9 \mathrm{cm}^{-2}$, the e-e scattering time is far longer than the LO phonon scattering time (~ 0.6 ps) and as such the depopulation of c_2 and c_3 is completely governed by LO phonon scattering. Similarly, LO phonon scattering from c_4 to c_1 (≈ 40 ps) is faster than e-e scattering below $n_{c4} \approx 10^{10} \mathrm{cm}^{-2}$, so even at very high generation levels e-e scattering only plays a marginal role. In fact, at electron temperatures exceeding 40K the parasitic 4-2 and 4-3 LO phonon channels
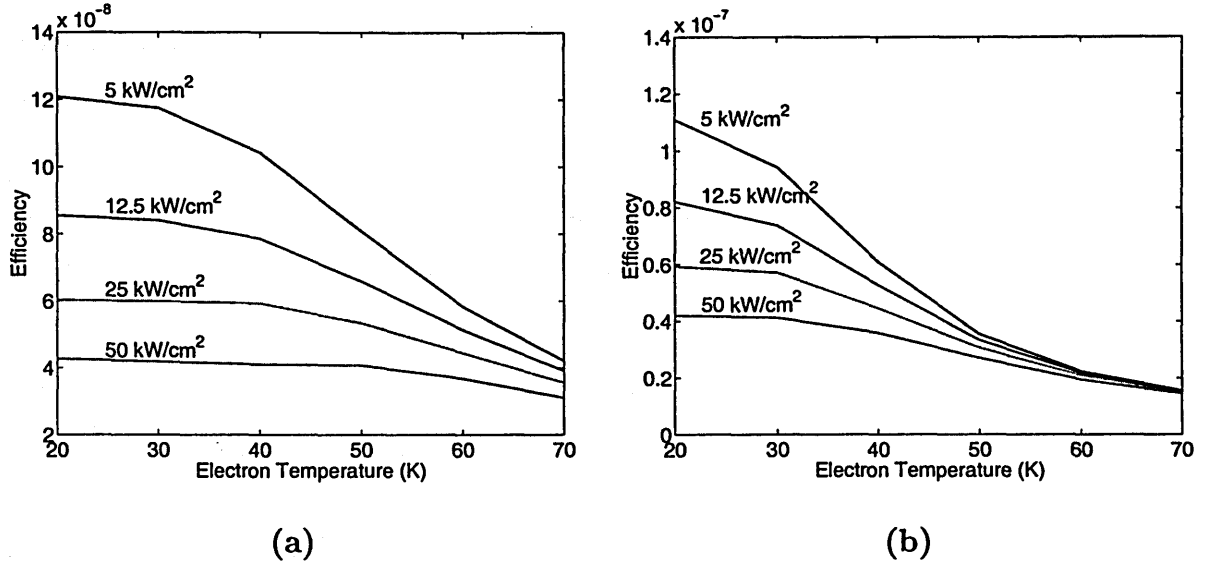
Figure 4-12: *Emission efficiency for the intersubband THz transition in the* (a)*three-and* (b)*four-level systems, for varying pump intensities.*

will start to dominate c_4 depopulation.

The lack of impurity scattering in the undoped structure promises a narrow spontaneous intersubband linewidth. In the simulations, a linewidth of 2 meV was used.

Results

A plot of the THz emission efficiency in both designs is shown in figure (4-12). At low temperatures, e-e scattering is the dominant depopulation mechanism in the upper intersubband laser level. The lower population densities generated by the lower-intensity pump beams have a longer e-e scattering time, therefore electrons are more likely to stay in the upper subband and make a radiative transition to c_3, increasing the emission efficiency. At high temperatures, LO phonon scattering is the main depopulation process. As LO phonon scattering times do not depend on the population density of the initial subband, $\tau_4^{tot} \approx \tau_4^{LO}(T)$ for all pump intensities at high temperatures.

In the 3-level system the LO phonon scattering channel from the upper THz laser level to the c_1 state is more pronounced than in the 4-level design. This results in a lower electron population in the upper intrasubband laser level, and a lower gain. As the temperature increases, more electrons in the high energy tail of the Fermi distribution of the upper CB level will have sufficient energy to emit an LO phonon and scatter to the lower laser level (figure (4-13)). The shortened lifetime in the upper state gives rise to a smaller population inversion, and hence a decrease in gain (figures (4-15), (4-17)). Eventually, the LO phonon scattering between the intersubband laser levels destroys the population inversion and the gain becomes negative.

The intersubband gain in the three-level design is far lower than in the four-level design (figure (4-18)). This is due to the insufficient difference in the (LO-phonon-mediated) depopulation of the upper and lower THz laser levels (figure (4-14)). Already at 50K the parasitic LO phonon channel 3-2 depopulates the upper state. The 2-1 scattering rate is too low to maintain the population inversion, but is difficult to increase without dramatically increasing the 3-1 LO phonon scattering rate in the process (figure (4-16)). In the four-level system the fast depopulation of the lower laser level virtually guarantees a population inversion, even as the parasitic LO phonon channel becomes more prominent with increasing temperature. Note that the short LO phonon scattering times for c_3 and c_2 to c_1 assume a coherent electron wave function in c_3 and c_2. This means that the thick barrier between the well containing c_1 and the other wells does not induce significant scattering and localization of c_2 and c_3.

In figure (4-19) the intersubband gain is plotted against the pump intensity for the four-level scheme. The approximate square-root dependence of the gain on the pump beam intensity is the result of the quasi-linear dependence of e-e scattering rates on population density. The population density in the upper THz laser level can be written as $n_{c4} = G_{c4}\tau_{c4}$. The scattering time τ_{c4} at low temperatures is mainly due to e-e scattering, and hence inversely proportional to n_{c4}. We conclude that

$n_{c4}^2 \propto G_{c4}$, or gain $\propto n_{c4} \propto G_{c4}^{1/2}$.

In the three-level system (figure (4-20)) the LO phonon depopulation lifetime from the lower THz laser level to c_1 is on the order of 10 ps, and from c_3 to c_1 32 ps as seen in figure (4-14). For the higher generation rates (and hence higher population densities), e-e scattering rates between the aforementioned levels can easily exceed the LO phonon scattering rates. At low temperatures, the e-e scattering time between the THz laser levels is the main depopulation mechanism for c_3. The small dip in the gain curve indicates where the $c_3 - c_1$ LO phonon scattering rate and the $c_3 - c_2$ e-e scattering rate become equal. For higher generation rates, the gain remains positive because the $c_2 - c_1$ electron-electron scattering rate keeps pace with the $c_3 - c_2$ rate (figure (4-21)).

Finally, it is interesting to ask the question whether the pump threshold level is lower for the intersubband or the interband laser. If the threshold is lower for the interband laser, then n_{c1} is clamped and a red light laser will coexist with the intersubband THz laser. If it is the other way around, then n_{c1} is not clamped and a red light LED will accompany the intersubband THz laser.

In figure (4-22) the threshold pump intensities for the interband and intersubband lasers in the four-level system are indicated for varying temperatures. The intersubband threshold displays a high temperature sensitivity, reflecting the opening of a LO phonon channel between the upper THz laser level and c_2 and c_3. Note that the interband threshold pump intensity is lower than the intersubband threshold, predicting an interband laser at pump power levels which allow intersubband lasing.

4.3.4 Simulation Conclusion

In the design of an optically pumped intersubband THz laser, there are many parameters to keep in mind. The proposed designs consist generally of three levels, two levels involved in the actual THz laser and one "collector" level. The lower THz laser level is depopulated by LO phonon scattering. In this design, as in any laser design,
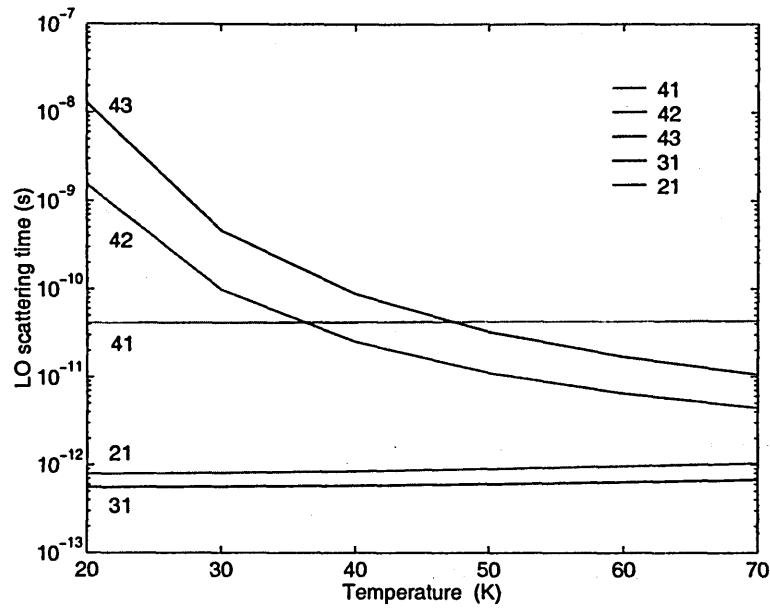
Figure 4-13: *LO phonon scattering times vs. temperature at a pump power density of 5000 W/cm² for the 4-level system.*
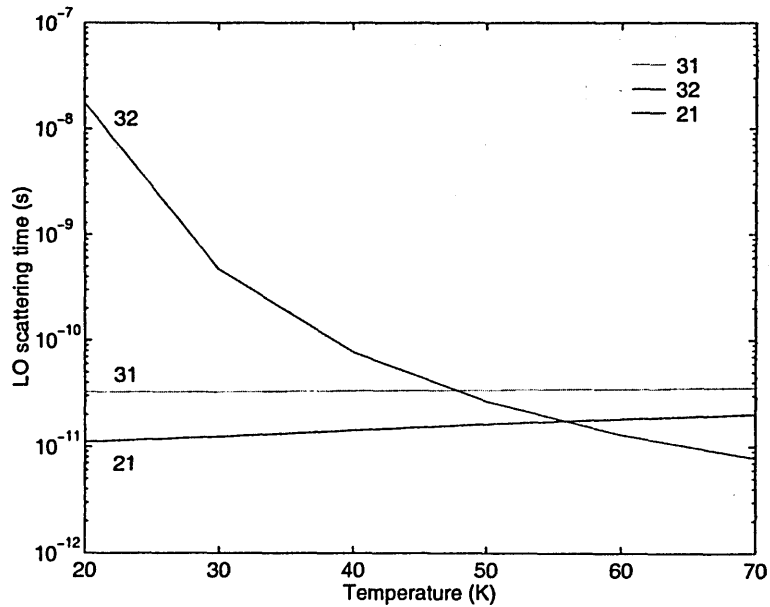
Figure 4-14: *LO phonon scattering times vs. temperature at a pump power density of 5000 W/cm² for the 3-level system.*
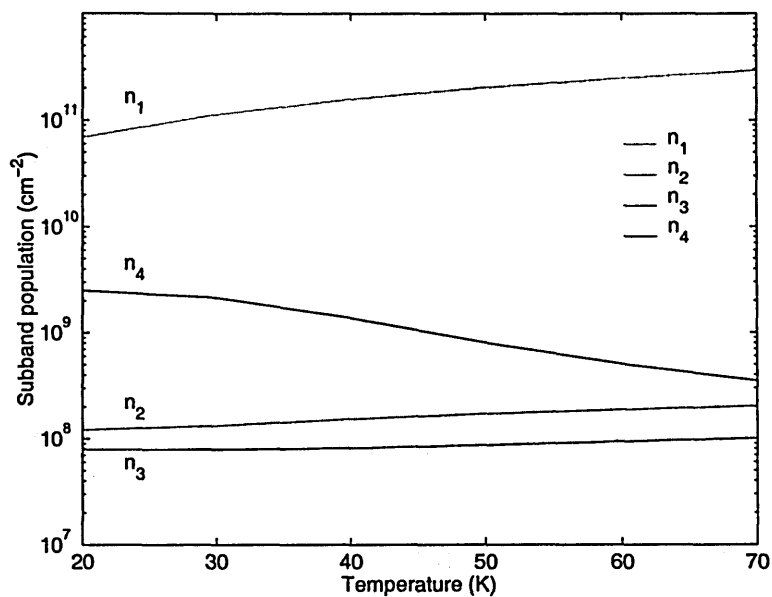
Figure 4-15: *Carrier density vs. temperature at pump power density of 5000 W/cm²*
for the 4-level system.
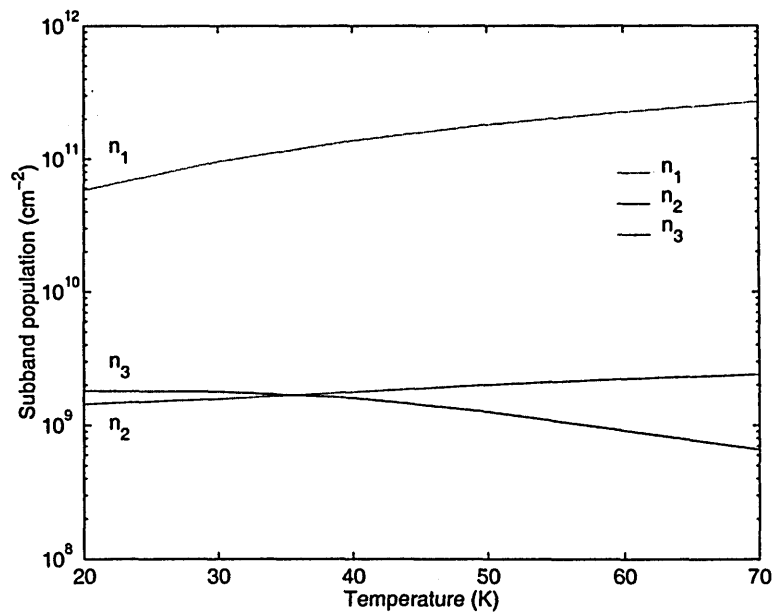
Figure 4-16: *Carrier density vs. temperature at a pump power density of 5000 W/cm²*
for the 3-level system.
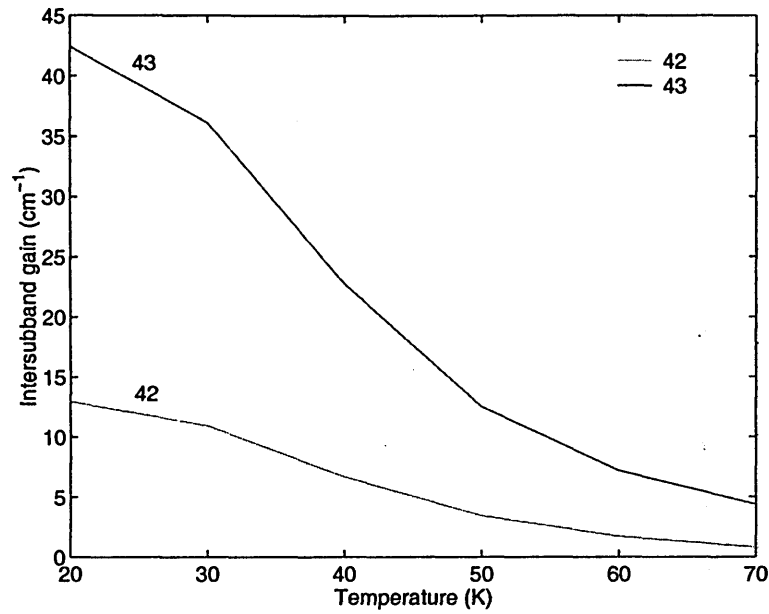
Figure 4-17: *Intersubband gain vs. temperature at a pump power density of 5000 W/cm^2 for the 4-level system, for a gain linewidth of 2 meV.*
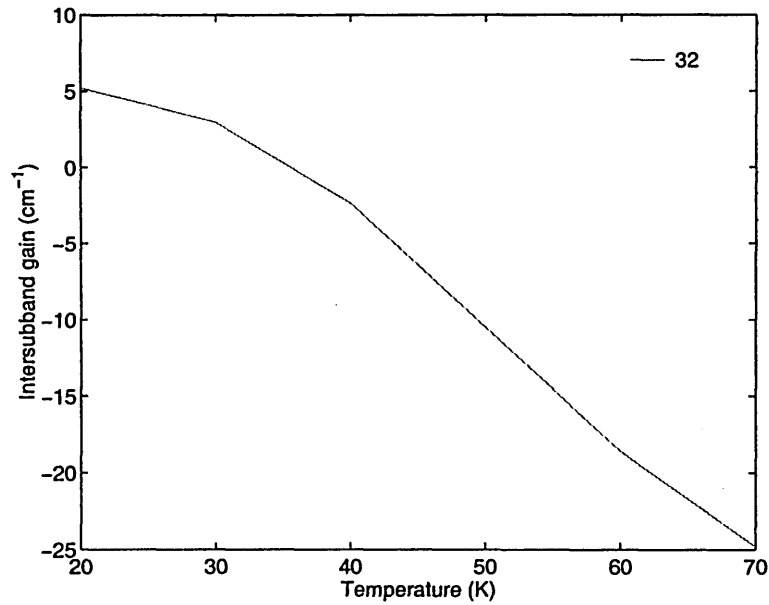
Figure 4-18: *Intersubband gain vs. temperature at a pump power density of 5000 W/cm^2 for the 3-level system, for a gain linewidth of 2 meV.*

a key issue is how to maximize carrier lifetime in the upper laser level while still depopulating the lower laser level as fast as possible. The three-level design tries to spatially separate the upper THz laser level and c_1, while the four-level design spatially extends the lower THz laser level to ensure faster depopulation and still keeps the aforementioned spatial separation.

A self-consistent calculation using the rate-equations governing the subband populations was performed, including carrier-carrier scattering, LO phonon scattering and radiative recombination and generation. The simulation of the proposed three- and four-level systems clearly indicates that the four-level system is the most promising. The fast depopulation of the lower THz laser levels is the key to the superior simulation results as compared to the three-level design. It is not possible to selectively pump the upper THz laser level, which makes it even more important to evacuate the lower THz laser level as efficiently as possible. The calculated gain (20 cm^{-1} at 50K) at pump beam intensities of 5-10 kW/cm^2 can suffice to obtain THz lasing action, provided that the cavity losses can be kept in check. The performance of the THz laser is predicted to be very dependent on electron temperature, mainly due to the opening of a parasitic LO-phonon channel between the THz laser levels. Interband lasing seems to be easier to obtain, as the calculated threshold pump intensity is lower than for the intersubband case.

The logical next step is manufacturing a device and building an experimental setup. The simulated results suggest that a 0.5 W pump source focused on a device with area 20$\times$500 μm^2 (corresponding to 5 kW/cm^2) could be sufficient to obtain THz lasing action. Doing this will enable us to check the model used and provide insight into several topics, such as electron temperature and cavity losses for interband and intersubband laser. We will be able to see how severe the free carrier absorption is, caused by the accumulated carriers in the lowest subband in each module.
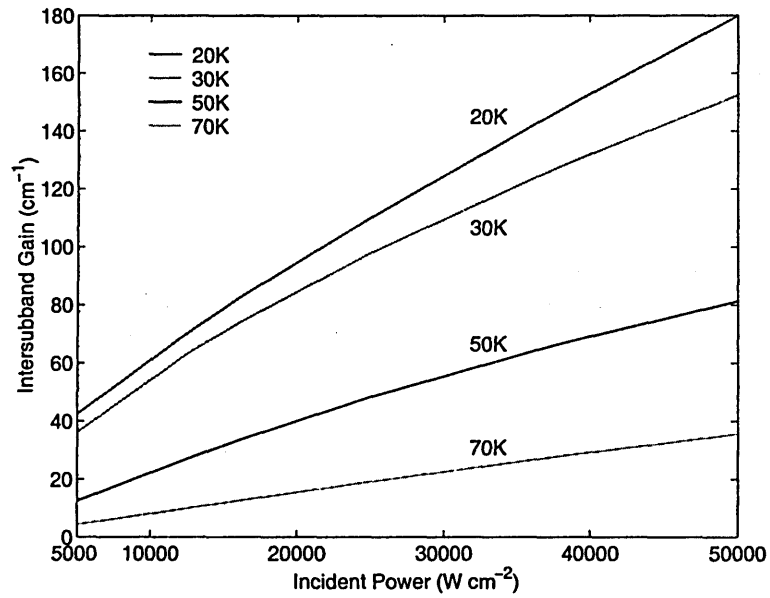
Figure 4-19: *Intersubband gain 4-3 as a function of pump intensity and various electron temperatures, for the four-level system. The intersubband spontaneous emission linewidth is 2 meV.*
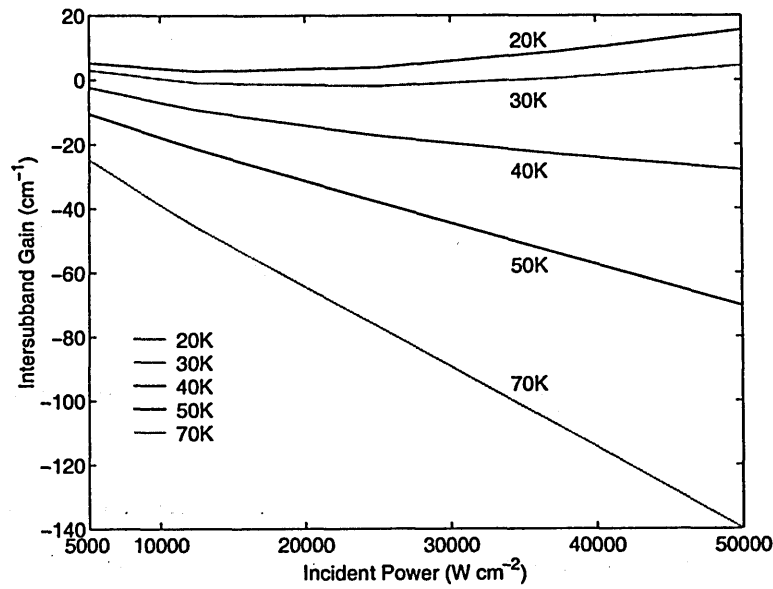
Figure 4-20: *Intersubband gain 3-2 as a function of pump intensity and various electron temperatures, for the three-level system. The intersubband spontaneous emission linewidth is 2 meV.*
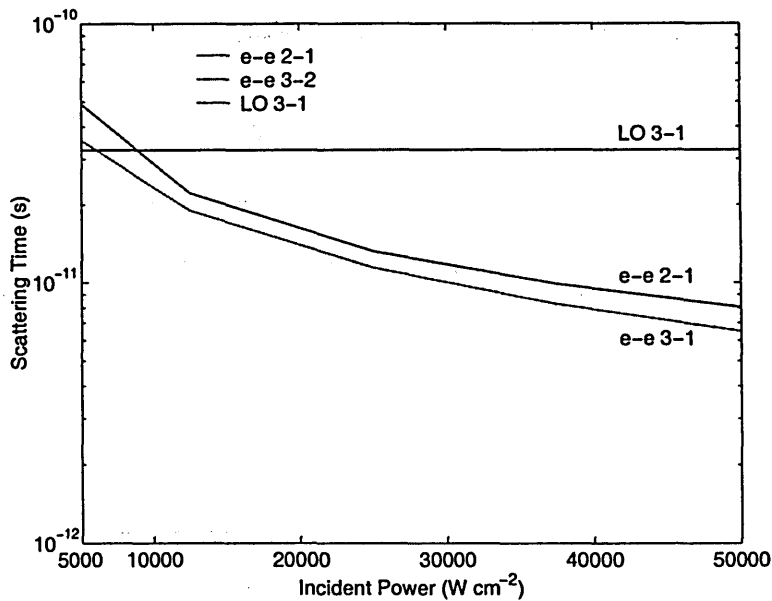
Figure 4-21: *Intersubband scattering times as a function of pump intensity for an electron temperature of 20 K, in the three-level system.*
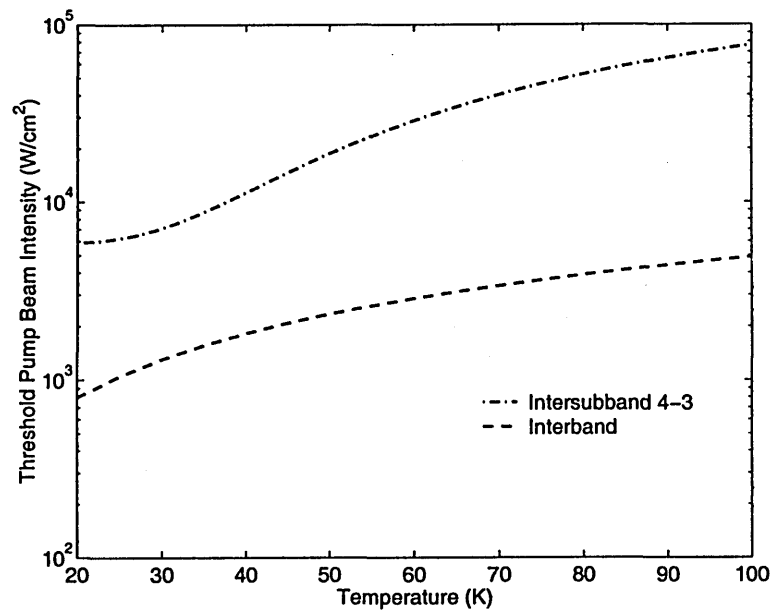
Figure 4-22: *Pump beam intensity thresholds for the interband and intersubband lasers for varying temperatures. The lasing threshold for the interband laser is always reached first. The gain threshold was taken to be 50 cm^{-1} in both cases.*

Appendix A

Appendix : Matlab Scripts and Functions

A.1 Finite Difference Method in the 2-Band Model

The method used to solve the coupled Hamiltonian equations in the two-band model is based on the shooting method as described by Paul Harrison for the simple case of the conduction band. [39]

As a starting point, we consider the general one-dimensional form of the Hamiltonian equations (see equations (2.53)-(2.56)). In order to allow for a variable effective mass, we rewrite these equations as:

$$
\begin{bmatrix} H_{hh} + V & W \\ W^\dagger & H_{lh} + V \end{bmatrix} \begin{bmatrix} F_{hh} \\ F_{lh} \end{bmatrix} = E(\mathbf{k}) \begin{bmatrix} F_{hh} \\ F_{lh} \end{bmatrix},
\tag{A.1}
$$

with

$$
\hat{H}_{lh} = -\frac{\partial}{\partial z}(\gamma_1 + 2\gamma_2)\frac{\partial}{\partial z} + (\gamma_1 - \gamma_2)k_t^2,
\tag{A.2}
$$

$$
\hat{H}_{hh} = -\frac{\partial}{\partial z}(\gamma_1 - 2\gamma_2)\frac{\partial}{\partial z} + (\gamma_1 + \gamma_2)k_t^2,
\tag{A.3}
$$

$$
\hat{W} = \sqrt{3}k_t(\gamma_2 k_t - 2\gamma_3\frac{\partial}{\partial z}).
\tag{A.4}
$$

The potential $V(z)$ describes the valence band edge of the quantum well structure (in terms of hole energy), and F_{hh} and F_{lh} represent the hole wavefunction while under the effective mass and envelope function approximations.

We can rewrite the effective mass equations as :

$$\begin{bmatrix} H_{hh} + V(z) - E & W \\ W^\dagger & H_{lh} + V(z) - E \end{bmatrix} \begin{bmatrix} F_{hh} \\ F_{lh} \end{bmatrix} = 0. \tag{A.5}$$

The problem now is to find a numerical method for the solution of both the energy eigenvalues E and the eigenfunctions F for any $V(z)$.

For this purpose, we can expand the first and second derivatives in terms of finite differences. e can approximate the first derivative of a function $f(z)$ as :

$$\frac{df}{dz} \approx \frac{\Delta f}{\Delta z} = \frac{f(z + \delta z) - f(z - \delta z)}{2\delta z}. \tag{A.6}$$

The second derivative follows as:

$$\frac{d^2 f}{dz^2} \approx \frac{\left.\frac{df}{dz}\right|_{z+\delta z} - \left.\frac{df}{dz}\right|_{z-\delta z}}{2\delta z}, \tag{A.7}$$

$$= \frac{f(z + 2\delta z) - 2f(z) + f(z - 2\delta z)}{(2\delta z)^2}. \tag{A.8}$$

As δz is an, as yet, undefined small step along the z-axis, and as it only appears in equation (A.8) with the factor 2, then we can simplify this expression by substituting δz for $2\delta z$, i.e. :

$$\frac{d^2 f}{dz^2} \approx \frac{f(z + \delta z) - 2f(z) + f(z - \delta z)}{(\delta z)^2}. \tag{A.9}$$

Let's take a closer look at the term $H_{lh}^0 = \frac{\partial}{\partial z}(\gamma_1 + 2\gamma_2)\frac{\partial F_{lh}}{\partial z}$ in the light hole Hamiltonian, and try to express this in terms of finite differences. We can rewrite this term as:

$$H_{lh}^0 = \frac{\partial}{\partial z}[\gamma_1 + \gamma_2]\frac{\partial F_{lh}}{\partial z} + (\gamma_1 + \gamma_2)\frac{\partial^2 F_{lh}}{\partial z^2}. \tag{A.10}$$

However, shooting equations derived from this point by expanding the derivatives in terms of finite differences have led to significant computational inaccuracies in systems with a large discontinuous change in the effective mass (the Luttinger parameters), as occurs in the GaAs/AlGaAs system. The source of the inaccuracies is thought to arise from the δ-function nature of the effective-mass derivative.

A more robust scheme can be derived by expanding H_{lh}^0 starting from the left-hand derivative :

$$H_{lh}^0 \approx \frac{(\gamma_1 + \gamma_2)\frac{\partial F_{lh}}{\partial z}\Big|_{z+\delta z} - (\gamma_1 + \gamma_2)\frac{\partial F_{lh}}{\partial z}\Big|_{z-\delta z}}{2\delta z} \tag{A.11}$$

Recalling the centered finite difference expansion for the first derivative, we can write the numerator of the above expression as :

$$2\delta z H_{lh}^0 = (\gamma_1 + 2\gamma_2)|_{z+\delta z} \frac{F_{lh}(z + 2\delta z) - F_{lh}(z)}{2\delta z}$$
$$- (\gamma_1 + 2\gamma_2)|_{z-\delta z} \frac{F_{lh}(z) - F_{lh}(z - 2\delta z)}{2\delta z}, \tag{A.12}$$

or

$$(2\delta z)^2 H_{lh}^0 = (\gamma_1 + 2\gamma_2)|_{z+\delta z} (F_{lh}(z + 2\delta z)$$
$$- F_{lh}(z)) - (\gamma_1 + 2\gamma_2)|_{z-\delta z} (F_{lh}(z) - F_{lh}(z - 2\delta z)). \tag{A.13}$$

Making the transformation $2\delta z \to \delta z$ then yields:

$$H_{lh}^0 = \frac{1}{(\delta z)^2} \left[(\gamma_1 - 2\gamma_2)^+ F_{lh}(z + \delta z) - \left[(\gamma_1 - 2\gamma_2)^+ + (\gamma_1 - 2\gamma_2)^-\right] F_{lh}(z) \right.$$
$$\left. + (\gamma_1 - 2\gamma_2)^- F_{lh}(z - \delta z) \right], \tag{A.14}$$

with

$$(\gamma_1 + 2\gamma_2)^+ = (\gamma_1 + 2\gamma_2)|_{z+\delta z/2}, \tag{A.15}$$

$$(\gamma_1 + 2\gamma_2)^- = (\gamma_1 + 2\gamma_2)|_{z-\delta z/2}, \tag{A.16}$$

$$(\gamma_1 - 2\gamma_2)^+ = (\gamma_1 - 2\gamma_2)|_{z+\delta z/2}, \tag{A.17}$$

$$(\gamma_1 - 2\gamma_2)^- = (\gamma_1 - 2\gamma_2)|_{z-\delta z/2}. \tag{A.18}$$

We now substitute the finite difference expressions for $\partial/\partial z$, H^0_{lh} and a similar expression for its heavy-hole counterpart into the effective mass equations, and obtain:

$$\begin{aligned}
0 &= -\frac{(\gamma_1 - 2\gamma_2)^+}{(\delta z)^2} F_{hh}(z + \delta z) + \frac{(\gamma_1 - 2\gamma_2)^+ + (\gamma_1 - 2\gamma_2)^-}{(\delta z)^2} F_{hh}(z) \\
&\quad -\frac{(\gamma_1 - 2\gamma_2)^-}{(\delta z)^2} F_{hh}(z - \delta z) + (\gamma_1 + \gamma_2) k_t^2 F_{hh}(z) \\
&\quad +(V(z) - E) F_{hh}(z) + \sqrt{3}\gamma_2 k_t^2 F_{lh}(z) \\
&\quad -2\sqrt{3}\gamma_3 k_t \frac{F_{lh}(z + \delta z) - F_{lh}(z - \delta z)}{2\delta z},
\end{aligned} \tag{A.19}$$

$$\begin{aligned}
0 &= \sqrt{3}\gamma_2 k_t^2 F_{hh}(z) + 2\sqrt{3}\gamma_3 k_t \frac{F_{hh}(z + \delta z) - F_{hh}(z - \delta z)}{2\delta z} \\
&\quad -\frac{(\gamma_1 + 2\gamma_2)^+}{(\delta z)^2} F_{lh}(z + \delta z) + \frac{(\gamma_1 + 2\gamma_2)^+ + (\gamma_1 + 2\gamma_2)^-}{(\delta z)^2} F_{lh}(z) \\
&\quad -\frac{(\gamma_1 + 2\gamma_2)^-}{(\delta z)^2} F_{lh}(z - \delta z) + (\gamma_1 - \gamma_2) k_t^2 F_{lh}(z) + (V(z) - E) F_{lh}(z). \tag{A.20}
\end{aligned}$$

The Luttinger parameters γ can be found at the intermediary points $z \pm \delta z/2$ by taking the mean of the two neighboring points at z and $z \pm \delta z$.

We see that we draw up a set of finite difference equations if we map the potential $V(z)$ and the Luttinger parameters γ to a grid along the z-axis. We will now try to solve these coupled equations for the eigenenergies E and the eigenfunctions F. Assuming an equidistant grid z_i, with a grid step δz, we can substitute $z \to z_i$, $z - \delta z \to z_{i-1}$ and $z + \delta z \to z_{i+1}$.

If we assume a given energy E, we still have 6 unknown parameters in the finite difference equations. However, we can rewrite these equations so that we are able to find $F_{lh}(z_{i+1})$ and $F_{hh}(z_{i+1})$ from their values at the two previous nodes, $F_{lh}(z_{i-1})$, $F_{lh}(z_-)$, $F_{hh}(z_{i-1})$, $F_{hh}(z_i)$.

$$F_{hh}(z_{i+1}) \left[1 + 3\frac{\gamma_3^2}{(\gamma_1 + 2\gamma_2)^+ (\gamma_1 - 2\gamma_2)^+} k_t^2 (\delta z)^2 \right]$$

$$= F_{hh}(z_{i-1}) \left[-\frac{(\gamma_1 - 2\gamma_2)^-}{(\gamma_1 - 2\gamma_2)^+} + 3\frac{\gamma_3^2}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+} k_t^2(\delta z)^2 \right]$$

$$+ F_{lh}(z_{i-1}) \left[\sqrt{3}\frac{\gamma_3}{(\gamma_1 - 2\gamma_2)^+} k_t \delta z (1 + \frac{(\gamma_1 + 2\gamma_2)^-}{(\gamma_1 + 2\gamma_2)^+}) \right]$$

$$+ F_{hh}(z_i) \left[\frac{(\gamma_1 - 2\gamma_2)^+ + (\gamma_1 - 2\gamma_2)^-}{(\gamma_1 - 2\gamma_2)^+} + \frac{\gamma_1 + \gamma_2}{(\gamma_1 - 2\gamma_2)^+} k_t^2(\delta z)^2 \right.$$
$$\left. + \frac{V(z_i) - E}{(\gamma_1 - 2\gamma_2)^+}(\delta z)^2 - 3\frac{\gamma_3 \gamma_2}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+} k_t^3(\delta z)^3 \right]$$

$$+ F_{lh}(z_i) \left[\sqrt{3}\frac{\gamma_2}{(\gamma_1 - 2\gamma_2)^+} k_t^2(\delta z)^2 \right.$$
$$- \sqrt{3}\frac{\gamma_3}{(\gamma_1 - 2\gamma_2)^+} \frac{(\gamma_1 + 2\gamma_2)^+ + (\gamma_1 + 2\gamma_2)^-}{(\gamma_1 + 2\gamma_2)^+} k_t \delta z$$
$$- \sqrt{3}\frac{\gamma_3(\gamma_1 - \gamma_2)}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+} k_t^3(\delta z)^3$$
$$\left. - \sqrt{3}\frac{\gamma_3}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+}(V(z_i) - E)k_t(\delta z)^3 \right] \tag{A.21}$$

$$F_{lh}(z_{i+1}) \left[1 + 3\frac{\gamma_3^2}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+} k_t^2(\delta z)^2 \right]$$

$$= F_{hh}(z_{i-1}) \left[\sqrt{3}\frac{\gamma_3}{(\gamma_1 + 2\gamma_2)^+} k_t \delta z \left(-1 - \frac{(\gamma_1 - 2\gamma_2)^-}{(\gamma_1 - 2\gamma_2)^+} \right) \right]$$

$$+ F_{lh}(z_{i-1}) \left[-\frac{(\gamma_1 + 2\gamma_2)^-}{(\gamma_1 + 2\gamma_2)^+} + 3\frac{\gamma_3^2}{(\gamma_1 + 2\gamma_2)^+(\gamma_1 - 2\gamma_2)^+} k_t^2(\delta z)^2 \right]$$

$$+ F_{hh}(z_i) \left[\sqrt{3}\frac{\gamma_2}{(\gamma_1 + 2\gamma_2)^+} k_t^2(\delta z)^2 \right.$$
$$\left. + \sqrt{3}\frac{\gamma_3}{(\gamma_1 + 2\gamma_2)^+} \frac{(\gamma_1 - 2\gamma_2)^+ + (\gamma_1 - 2\gamma_2)^-}{(\gamma_1 - 2\gamma_2)^+} k_t \delta z \right.$$

$$+\sqrt{3}\frac{\gamma_3(\gamma_1+\gamma_2)}{(\gamma_1+2\gamma_2)^+(\gamma_1-2\gamma_2)^+}k_t^3(\delta z)^3$$

$$+\sqrt{3}\frac{\gamma_3}{(\gamma_1+2\gamma_2)^+(\gamma_1-2\gamma_2)^+}(V(z_i)-E)k_t(\delta z)^3\Bigg]$$

$$+ F_{lh}(z_i)\left[\frac{(\gamma_1+2\gamma_2)^++(\gamma_1+2\gamma_2)^-}{(\gamma_1+2\gamma_2)^+}+\frac{\gamma_1-\gamma_2}{(\gamma_1+2\gamma_2)^+}k_t^2(\delta z)^2\right.$$

$$\left.+\frac{V(z_i)-E}{(\gamma_1+2\gamma_2)^+}(\delta z)^2+3\frac{\gamma_2\gamma_3}{(\gamma_1+2\gamma_2)^+(\gamma_1-2\gamma_2)^+}k_t^3(\delta z)^3\right] \quad \text{(A.22)}$$

These equations imply that, if the wavefunctions are known at the two points $(z-\delta z)$ and z, then the value of the wavefunctions at $(z+\delta z)$ can be determined for any energy E. This iterative equation forms the basis of a standard method of solving equations numerically, and is known as *the shooting method.*

The equations can be rewritten in a matrix formalism, which allows for an easy implementation in a MATLAB program code. If we use use a coefficient notation for equations (A.21) and (A.22), i.e.

$$F_{hh}(z_{i+1}) = a_1 F_{hh}(z_{i-1} + a_2 F_{lh}(z_{i-1}) + a_3 F_{hh}(z) + a_4 F_{hh}(z), \quad \text{(A.23)}$$

$$F_{lh}(z_{i+1}) = b_1 F_{hh}(z_{i-1} + b_2 F_{lh}(z_{i-1}) + b_3 F_{hh}(z) + b_4 F_{hh}(z), \quad \text{(A.24)}$$

the effective mass equations can be written in a recursive "transfer" matrix expression :

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}\begin{bmatrix} F_{hh}(z_{i-1}) \\ F_{lh}(z_{i-1}) \\ F_{hh}(z_i) \\ F_{lh}(z_i) \end{bmatrix} = \begin{bmatrix} F_{hh}(z_i) \\ F_{lh}(z_i) \\ F_{hh}(z_{i+1}) \\ F_{lh}(z_{i+1}) \end{bmatrix}. \quad \text{(A.25)}$$

Provided that we have initial values for the wave functions at the first and second node, we can determine the wavefunction values at any node by an iterative procedure.

By multiplying matrices, it is possible to obtain an expression for the wavefunction values at any node in function of the initial values:

$$
\begin{bmatrix}
F_{hh}(z_n) \\
F_{lh}(z_n) \\
F_{hh}(z_{n+1}) \\
F_{lh}(z_{n+1})
\end{bmatrix}
= \mathbf{M}_{n+1}\mathbf{M}_n\mathbf{M}_{n-1}\cdots\mathbf{M}_3\mathbf{M}_2
\begin{bmatrix}
F_{hh}(z_0) \\
F_{lh}(z_0) \\
F_{hh}(z_1) \\
F_{lh}(z_1)
\end{bmatrix}.
\tag{A.26}
$$

Now the questions remain what a suitable choice is for those initial values, and how to determine whether an energy is an eigenenergy or not.

A.2 Initial Conditions and Practical Implementation

Using four known values of the wavefunction components at z and $(z + \delta z)$, a fifth and sixth value can be predicted. Using this new point together with the known wavefunction components at z, we can subsequently find the wavefunctions at $(z + 2\delta z)$, and so on. Hence the complete wave function solution can be found for any particular energy. The solutions for steady states have wavefunctions which satisfy the standard boundary conditions, i.e. :

$$
F \to 0 \text{ and } \frac{\partial}{\partial z}F \to 0, \text{ as } z \to \pm\infty
\tag{A.27}
$$

As argued by Harrison [39], in the one-band case of the conduction band only two initial values are required, and a suitable choice is "0" for the first node, and "1" for the second point. The "1" can be any arbitrary number, as changing it will only scale the wavefunction (the finite difference equations are linear) and this does not affect the eigenenergy. The valence band case is a bit more complicated, as now there are two coupled wavefunction components, and one cannot be scaled independently from the other. Therefore we choose the initial values to be "0" and "1" for one band, and

"0" and "c" for the other band. Here c is a parameter, which is to be determined when solving the coupled equations.

The energy is varied systematically until both the wavefunction components switch from diverging to $\pm\infty$ to $\mp\infty$, satisfying the boundary conditions. However, an additional problem is that the extra parameter c complicates matters. On top of that, in many cases one of the wavefunction components exhibit very sharp sign switches, often a component switches signs twice within one energy search step. In order to work around these problems, we minimize the amplitude of the wavefunction at the end of the grid. The function to be minimized can be found by generating the transfer matrix which propagates the wavefunction at the first two nodes to the last two notes:

$$
\begin{bmatrix} F_{hh}(z_{N-1}) \\ F_{lh}(z_{N-1}) \\ F_{hh}(z_N) \\ F_{lh}(z_N) \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ c \end{bmatrix}. \tag{A.28}
$$

We then minimize the wavefunction amplitude at the final node :

$$
(m_{33} + m_{34}c)^2 + (m_{43} + m_{44}c)^2 \rightarrow c_{min}. \tag{A.29}
$$

We then look for a minimum in $c_{min}(E)$. A solution for the Hamiltonian equations is found when this minimum wavefunction amplitude is smaller than a certain threshold value. This guarantees a converging wavefunction. The wavefunction can be found by substituting c_{min} for c in equation (A.26).

The wavefunctions obtained from this numerical method are not normalized. This can easily be achieved with the following transformation

$$
\begin{bmatrix} F_{hh}(z) \\ F_{lh}(z) \end{bmatrix} \rightarrow \begin{bmatrix} F_{hh}(z) \\ F_{lh}(z) \end{bmatrix} \frac{1}{\sqrt{\int (F_{hh}^2(z) + F_{lh}^2(z))dz}} \tag{A.30}
$$

A.3 Code

lutt.m

This function looks for the file containing the Al content x.r, as generated by Paul Harrison's code. It then calculates and saves the three Luttinger parameters for the structure specified in x.r, and saves them as ascii .r-files in g1.r, g2.r and g3.r. The effective mass data was taken from Adachi [26].

```
function lutt;
%Create files containing Luttinger parameters g1, g2, g3 vs. z
%Input is x.r

xr=load('x.r');          % Al content in AlGaAs
z=xr(:,1);               % z grid
x=xr(:,2)';              % Al content

m0=9.1e-31;              % e mass
hbar=1.05e-34;
mhh=0.50+0.29*x;         % heavy hole mass
mlh=0.087+0.063*x;       % light hole mass
mhh110=0.85*mhh./mhh;    % hh mass along <110>

%Luttinger parameters gamma for AlGaAs
coeff=[1 -2 0 ; 1 2 0 ; 1 0 -2];
const=hbar^2/(2*m0) *[1./mhh ; 1./mlh; 1./mhh110];
g=coeff^-1*const;
g=g';

g1=[z g(:,1)];
g2=[z g(:,2)];
g3=[z g(:,3)];
save 'g1.r' g1 -ascii -double
save 'g2.r' g2 -ascii -double
save 'g3.r' g3 -ascii -double
```

input_deck.m

Some input values for the dispersion relation calculator vbdispersion.m and the simple band-edge calculation script finelvb.m. This allows for easier manipulation of some key input parameters.

```
%finelvb and vbdispersion

min_en_diff=0.05*meV;    % used to determine whether two solutions are the same
alimit=1e5;              % maximum for a
kt=1e5;                  % kt for 'band edge' calculation
Estep=0.1*meV;           % E step for band edge calculation
Eacc=meV*1e-9;           % E resolution of calculated subband energies

%vbdispersion - search parameters

subband_en_search=0.15*meV;  % energy search range
en_abort=meV;                % abort search when E range > en_abort
ktmax=1e9;                   % maximum calculated kt
ktstep=2e6;                  % kt step for E-k dispersion
number_of_cb_subbands=4;     % # of cb subbands taken into account
%number_of_states=nr-1;       % # of vb subbands taken into account
```

vbdispersion.m

This is the main MATLAB script file, calculating the $E - k$ relations for all valence subbands. The function lutt.m is run first, to calculate the Luttinger parameters

for use in the transfer matrix code. The script works as explained in the theoretical section on finite difference calculations, searching an energy grid ranging over the energies allowing a bound solution. First, the zone center ($k = 0$) subband energies are determined by searching this energy grid and refining the energy solutions up to a certain energy accuracy E_{acc}. Two plots of the valence band structure are shown, distinguishing between heavy hole and light hole components of the wavefunctions. The wavefunction data is saved as an ascii file under the name wf_*subband_number*.dat.

The user can specify for how many valence subbands a dispersion relation is calculated, and at what resolution, by setting the corresponding parameters in input_deck.m. The dispersion relation is found by starting from the zone center energy, and increasing k to find a new value $E(k)$. The previous values for $E(k)$ are then used to predict a value which is used as an initial guess in the next k-step.

Simultaneously, the wavefunctions for the solutions are calculated and used to find the relative interband optical transition strength between each calculated valence subband and a user-specified number of conduction subbands. The relative transition strength data is saved as .mat-files.

```
%Shooting Method
clear
lutt

global z hbar meV V dz N g1 g2 g3

hbar=1.05e-34;
meV=1.602e-22;

%XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
% structural parameters and data
%XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

xr=load('x.r');
x=xr(:,2);            %Al content
z=xr(:,1);            %grid
Vr=load('v.r');       %potential profile
V=Vr(:,2);
g1r=load('g1.r');     %Luttinger parameters
g1=g1r(:,2);
g2r=load('g2.r');
g2=g2r(:,2);
g3r=load('g3.r');
g3=g3r(:,2);

input_deck           %input parameters

dz=z(2)-z(1);        %grid step size
N=size(V,1);         %number of grid points

cl=['m ';'c ';'r ';'g ';'b ';'y ';'m-.';'c-.';'r-.';'g-.';'b-.';'y-.';'m ';
'c ';'r ';'g ';'b ';'y ';'m-.';'c-.';'r-.';'g-.';'b-.';'y-.';'m ';'c ';'r ';'g ';'b ';'y '];
ch=['m ';'c ';'r ';'g ';'b ';'y ';'m-.';'c-.';'r-.';'g-.';'b-.';'y-.';'m ';
'c ';'r ';'g ';'b ';'y ';'m-.';'c-.';'r-.';'g-.';'b-.';'y-.';'m ';'c ';'r ';'g ';'b ';'y '];
nrm=1e7;

Emin=min(V);         %energy minimum
Emax=min(V(1),V(N)); %energy maximum

%XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
% calculation of band-edge subband energies and wavefunctions
%XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

mpl=figure;
plot(z/1e-10,V/meV,'k');
```

```
hold;
mph=figure;
plot(z/1e-10,V/meV,'k');
hold;

nr=1;
oldsignl=1;
oldsignh=1;
oldsignd=1;
oldsignd2=1;
oldsignd3=1;
index=1;
for E=Emin:Estep:Emax
   tfm=tfmatrix(E,kt);
   h10(index)=tfm(3,3);
   l10(index)=tfm(4,3);
   h01(index)=tfm(3,4);
   l01(index)=tfm(4,4);
   cmin(index)=-(tfm(1,3)*tfm(1,4)+tfm(2,3)*tfm(2,4))/(tfm(1,4)^2+tfm(2,4)^2);
   diffhmin(index)=tfm(3,3)+cmin(index)*tfm(3,4);
   difflmin(index)=tfm(4,3)+cmin(index)*tfm(4,4);
   dmin(index)=diffhmin(index)^2+difflmin(index)^2;
   dum(index)=abs(h10(index)*h01(index)+l01(index)*l10(index));

   if index>1
      diffdmin(index)=dmin(index)-dmin(index-1);
      newsignd=sign(diffdmin(index));
      if index>2
         if (newsignd-oldsignd==2)
E/meV
[a,Eresult]=shootvbd(E-2*Estep,E,kt,Eacc);
eval(['!echo diff  E = ' num2str(Eresult/meV,12)]);
if (nr==1) | ( (nr>1) & abs(Eresult-Estate(nr-1))>min_en_diff )
   Estate(nr)=Eresult;
   [fh,fl]=plotwav(Eresult,kt,1,a);
   if ( (fl(N)^2+fh(N)^2)/(max(abs(fl))^2+max(abs(fh))^2)<1e-4)
      Estate(nr)=Eresult;
      coeff(nr)=a;
      wfl(nr,:)=fl;
      wfh(nr,:)=fh;
      figure(mpl);
      plot(z/1e-10,Estate(nr)/meV+wfl(nr,:).^2/nrm,cl(nr,:));
      drawnow;
      refresh(mpl);
      figure(mph);
      plot(z/1e-10,Estate(nr)/meV+wfh(nr,:).^2/nrm,ch(nr,:));
      drawnow;
      refresh(mph);
      wf_save=[z fh' fl'];
      eval(['save wf_' num2str(nr) '.dat wf_save -ascii']);
      nr=nr+1;
   end
end
      end
      end
      oldsignd=newsignd;
   end

   if index>1
      diff3(index)=dum(index)-dum(index-1);
      newsignd3=sign(diff3(index));
      if index>2
         if (newsignd3-oldsignd3==2)
E/meV
[a,Eresult]=shootvbd(E-2*Estep,E+2*Estep,kt,Eacc,Estep/50);
eval(['!echo try  E = ' num2str(Eresult/meV,12)]);
if (nr==1) | ( (nr>1) & abs(Eresult-Estate(nr-1))>min_en_diff )
   Estate(nr)=Eresult;
   [fh,fl]=plotwav(Eresult,kt,1,a);
   if ( (fl(N)^2+fh(N)^2)/(max(abs(fl))^2+max(abs(fh))^2)<1e-4)
      Estate(nr)=Eresult;
      wfl(nr,:)=fl;
      wfh(nr,:)=fh;
      coeff(nr)=a;
      figure(mpl);
      plot(z/1e-10,Estate(nr)/meV+wfl(nr,:).^2/nrm,cl(nr,:));
      drawnow;
      refresh(mpl);
      figure(mph);
      plot(z/1e-10,Estate(nr)/meV+wfh(nr,:).^2/nrm,ch(nr,:));
      drawnow;
      refresh(mph);
      wf_save=[z fh' fl'];
      eval(['save wf_' num2str(nr) '.dat wf_save -ascii']);
      nr=nr+1;
   else
      [a,Eresult]=shootvbd(Eresult,E+2*Estep,kt,Eacc,Estep/50,2);
      eval(['!echo try  E = ' num2str(Eresult/meV,12)]);
      Estate(nr)=Eresult;
      [fh,fl]=plotwav(Eresult,kt,1,a);
      if ( (fl(N)^2+fh(N)^2)/(max(abs(fl))^2+max(abs(fh))^2)<1e-4)
         Estate(nr)=Eresult;
         wfl(nr,:)=fl;
         wfh(nr,:)=fh;
         figure(mpl);
         plot(z/1e-10,Estate(nr)/meV+wfl(nr,:).^2/nrm,cl(nr,:));
         drawnow;
         refresh(mpl);
         figure(mph);
         plot(z/1e-10,Estate(nr)/meV+wfh(nr,:).^2/nrm,ch(nr,:));
         drawnow;
         refresh(mph);
         wf_save=[z fh' fl'];
         eval(['save wf_' num2str(nr) '.dat wf_save -ascii']);
         nr=nr+1;
      end
```

```
    end
end
      end
    end
    oldsignd3=newsignd3;
  end

  Eindex(index)=E;
  index=index+1;
end

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
% calculation of E-k dispersion relations
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

clear wavcon rtsTE rtsTM

if ~exist('number_of_states')
  number_of_states=nr-1;
end

for cb_num=1:number_of_cb_subbands
  eval(['vf=load(''wf_e' num2str(cb_num) '.r'');']);
  wavcon(:,cb_num)=wf(:,2);
end

for nr=1:number_of_states

  clear subband_energy coeff

  E0=Estate(nr);                    % subband edge energy
  subband_energy(1)=E0;
  sbnd_en_srch_act=subband_en_search;   % search range
  sbnd_en_old=sbnd_en_srch_act;     % old search range (previous k)
  Estep=subband_en_search/10;       % energy step in range
  Estep_old=Estep;                  % old energy step
  del_en=0;                         % E difference between two previous E
  kt_index=1;
  abort=0;
  error_nf=0;                       % 1 if no minimum found
  sbnd_en_lim=meV*3e-6;             % minimum search range
  old_error=0;                      % diff between predicted and calculated E
  accuracy_modifier=1;
  num=1;                            % number of minima in search range
  tic
  Eacc_act=Eacc;
  ktmat=[(0:0.25:1.75)  2:(ktmax/ktstep)]*ktstep;

  while (kt_index<size(ktmat,2)) & ~abort    % kt iteration

    kt=ktmat(kt_index+1);
    min_en=subband_energy(kt_index)+del_en-sbnd_en_srch_act;
    max_en=subband_energy(kt_index)+del_en+sbnd_en_srch_act;
    min_en_arc(kt_index+1)=min_en;
    max_en_arc(kt_index+1)=max_en;
    [a,Eresult,error_nf]=shootvbd(min_en,max_en,kt,Eacc_act,Estep,num);

    if kt_index/50==floor(kt_index/50)
      drawnow
    end

XXXXXXXXXXXXXXXXXXX .Minimum not found  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

    if error_nf

      if (num==2) & (kt_index<4)
num=1;
      else

num_steps=2*sbnd_en_srch_act/Estep;

if (Estep>Estep_old/10) & (sbnd_en_srch_act<sbnd_en_srch_old*10)
  if num_steps<3e2
    Estep=Estep/2;
    num_steps=sbnd_en_srch_act/Estep;
    eval(['!echo Adjusting Estep to ' num2str(Estep/meV) ', ' num2str(floor(num_steps)) ' steps']);
  else
    sbnd_en_srch_act=sbnd_en_srch_act*2;
    Estep=sbnd_en_srch_act/10;
    eval(['!echo Too many steps - increasing search area to ' num2str(sbnd_en_srch_act/meV) ', Estep = ' num2str(Estep/meV) ', 10 steps']);
  end

elseif sbnd_en_srch_act<5*sbnd_en_srch_old
  sbnd_en_srch_act=sbnd_en_srch_old*5;
  Estep=Estep_old/2;
  num_steps=sbnd_en_srch_act/Estep;
  eval(['!echo Search area too small - increasing search area to ' num2str(sbnd_en_srch_act/meV) '
, Estep = ' num2str(Estep/meV) ', ' num2str(num_steps) ' steps']);

else
  ktmat_size=size(ktmat,2);
  ktstep_local=kt-ktmat(kt_index);
  kt_inter=kt-ktstep_local/2;
  ktmat=[ktmat(1:kt_index) kt_inter ktmat(kt_index+1:ktmat_size)];
  P=polyfit(ktmat((kt_index-3):kt_index)/ktstep,subband_energy((kt_index-3):kt_index)/meV,3);
  en_predict=polyval(P,ktmat(kt_index+1)/ktstep)*meV;
  del_en=en_predict-subband_energy(kt_index);
  sbnd_en_srch_act=sbnd_en_srch_old;
  Estep=Estep_old;
  eval(['!echo Inserting additional kt = ' num2str(kt)]);
end

if sbnd_en_srch_act>0.25*meV
  ktmat_size=size(ktmat,2);
  ktstep_local=kt-ktmat(kt_index);
  kt_inter=kt-ktstep_local/2;
```

```
    ktmat=[ktmat(1:kt_index) kt_inter ktmat(kt_index+1:ktmat_size)];
    P=polyfit(ktmat((kt_index-3):kt_index)/ktstep,subband_energy((kt_index-3):kt_index)/meV,3);
    en_predict=polyval(P,ktmat(kt_index+1)/ktstep)*meV;
    del_en=en_predict-subband_energy(kt_index);
    sbnd_en_srch_act=sbnd_en_srch_old;
    Estep=Estep_old;
    eval(['!echo Search area too large, inserting additional kt = ' num2str(kt)]);
end

error_nf=0;

    end

  else

XXXXXXXXXXXXXXXXXXXX  Minimum found  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

    [fh,fl]=plotwav(Eresult,kt,1,a);

    if ( ((fl(N)^2+fh(N)^2)/(max(abs(fl))^2+max(abs(fh))^2))<1e-4*accuracy_modifier )

subband_energy(kt_index+1)=Eresult;
coeff(kt_index+1)=a;

if kt_index<4
    del_en=(Eresult-subband_energy(kt_index))*(ktmat(kt_index+2)-kt)/(kt-ktmat(kt_index));
elseif kt_index<(size(ktmat,2)-1)
    P=polyfit(ktmat((kt_index-2):(kt_index+1))/ktstep,subband_energy((kt_index-2):(kt_index+1))/meV,3);
    en_predict=polyval(P,ktmat(kt_index+2)/ktstep)*meV;
    del_en=en_predict-Eresult;
end

for cb_num=1:number_of_cb_subbands
        ol1=trapz(z,(wavcon(:,cb_num).*fh'));
        ol2=trapz(z,(wavcon(:,cb_num).*fl'));
        rtsTE(nr,cb_num,kt_index+1)=0.5*(abs(ol1^2)+(1/3)*abs(ol2^2));
        rtsTM(nr,cb_num,kt_index+1)=(2/3)*abs(ol2^2);
end

old_error=Eresult-min_en-sbnd_en_srch_act;
kt_index=kt_index+1;

eval(['!echo kt=' num2str(kt) ', Eacc=' num2str(Eacc_act/meV) ' '
srch_delE='num2str(sbnd_en_srch_act/meV) ' Estep='num2str(Estep/meV) ' E found ' num2str(Eresult/meV) ' meV']);

sbnd_en_srch_old=sbnd_en_srch_act;
sbnd_en_srch_act=min([max(sbnd_en_srch_act/5, abs(5*(old_error))), subband_en_search]);
sbnd_en_srch_fnd=sbnd_en_srch_act;
if sbnd_en_srch_act<sbnd_en_lim;
    sbnd_en_srch_act=sbnd_en_lim;
end
sbnd_en_lim=max([sbnd_en_lim/10, abs(5*old_error), min(Eacc,Eacc_act*1e4)]);

Estep_old=Estep;
Estep=sbnd_en_srch_act/10;
Estep_fnd=Estep;

Eacc_prev=Eacc_act;
Eacc_act=min(Eacc*1e5,Eacc_act*1.25);
Eacc_fnd=Eacc_act;

accuracy_modifier=max(accuracy_modifier*.75,1);
changed_num=0;

        else

!echo Adjusting search parameters

if kt_index<3

  Eacc_act=Eacc_act/10;

  if Eacc_act<Eacc/20
    if num==1
      num=2;
    else
      num=1;
    end
  end
else
  Eacc_act=Eacc_act*.5;
  eval(['!echo Eacc changed to ' num2str(Eacc_act/meV)]);
  accuracy_modifier=min(accuracy_modifier*2,100);
  eval(['!echo accuracy modifier = ' num2str(accuracy_modifier)]);

  if Eacc_act<Eacc_prev/3
    ktmat_size=size(ktmat,2);
    ktstep_local=kt-ktmat(kt_index);
    kt_inter=kt-ktstep_local/2;
    ktmat=[ktmat(1:kt_index) kt_inter ktmat(kt_index+1:ktmat_size)];
    if kt_index>3
      P=polyfit(ktmat((kt_index-3):kt_index)/ktstep,subband_energy((kt_index-3):kt_index)/meV,3);
      en_predict=polyval(P,ktmat(kt_index+1)/ktstep)*meV;
    end
    del_en=en_predict-subband_energy(kt_index);
    sbnd_en_srch_act=sbnd_en_srch_old;
    Estep=Estep_old;
    eval(['!echo Inserting additional kt = ' num2str(kt_inter)]);
  end

  if Eacc_act<Eresult*eps*10
    if ~changed_num
      if num==1
num=2;
      else
num=1;
```

```
        end
      Estep=Estep_fnd;
      sbnd_en_srch_act=sbnd_en_srch_fnd;
      Eacc_act=Eacc_fnd;
      changed_num=1;
    else
      subband_energy(kt_index+1)=Eresult;
      eval(['!echo Solution not found, assumed calculated value']);

      old_error=Eresult-min_en-sbnd_en_srch_act;
      kt_index=kt_index+1;

      sbnd_en_srch_old=sbnd_en_srch_act;
      sbnd_en_srch_act=min([max(sbnd_en_srch_act, abs(5*(old_error))), subband_en_search]);
      if sbnd_en_srch_act<sbnd_en_lim;
sbnd_en_srch_act=sbnd_en_lim;
      end
      sbnd_en_lim=max([sbnd_en_lim/10, abs(5*old_error), min(Eacc,Eacc_act*1e4)]);

      Estep_old=Estep;
      Estep=sbnd_en_srch_act/10;

      Eacc_prev=Eacc_act;
      Eacc_act=Eacc_act*1e2;

      [fl,fh]=plotwav(Eresult,kt,1,a);
      figure
      plot(fh)
      hold;
      plot(fl,'r')
      drawnow

      a
      eval(['!echo ' num2str((fl(N)^2+fh(N)^2)/(max(abs(fl))^2+max(abs(fh))^2))]);

      accuracy_modifier=min(accuracy_modifier*2,100);
      eval(['!echo accuracy modifier = ' num2str(accuracy_modifier)]);
    end
  end

end
      end
    end
  end
t(nr)=toc;
eval(['!echo Elapsed time = ' num2str(t(nr)) ' s']);
figure
plot(ktmat,subband_energy/meV)
drawnow;
eval(['save En' num2str(nr) ' subband_energy']);
eval(['save kt' num2str(nr) ' ktmat']);


figure
plot(ktmat,squeeze(rtsTE(nr,1,1:size(ktmat,2))),cl(1));
hold
for cb_num=2:number_of_cb_subbands
    plot(ktmat,squeeze(rtsTE(nr,cb_num,1:size(ktmat,2))),cl(cb_num));
end
figure
plot(ktmat,squeeze(rtsTM(nr,1,1:size(ktmat,2))),cl(1)); .; .
hold
for cb_num=2:number_of_cb_subbands
    plot(ktmat,squeeze(rtsTM(nr,cb_num,1:size(ktmat,2))),cl(cb_num));
end

eval(['save coeff' num2str(nr) ' coeff'])

end

eval(['save rtsTE rtsTE'])
eval(['save rtsTM rtsTM'])
```

tfmatrix.m

The function `tfmatrix.m` returns the transfer matrix for the whole structure. Inputs are E, the energy in J, and kt, the in-plane hole momentum. This function implements the finite-difference equations (A.21) and (A.22).

```
function TFtot=tfmatrix(E,kt)
%function TFtot=tfmatrix(E,kt)
%returns the transfer matrix for E and kt
%E = subband energy (J)
%kt = in-plane hole momentum

if ~exist('kt')
  kt=0;
end
```

```
global g1 g2 g3 V N dz

TFtot=eye(4);
TF=zeros(4,4);
TF(1,3)=1;
TF(2,4)=1;

for n=2:(N-1)

  if (g1(n)~=g1(n-1)) | (n==2) %determine transfer matrix elements
    C1=g1(n)^2-4*g2(n)^2;

    Cf = 1 + 3*kt^2*dz^2*g3(n)^2/C1;

    TF(3,1) = (-1 + 3*kt^2*dz^2*g3(n)^2/C1) / Cf;

    TF(3,2) = (2*sqrt(3)*kt*dz*g3(n)/(g1(n)-2*g2(n))) / Cf;

    TF33 = ( 2 + kt^2*dz^2*(g1(n)+g2(n))/(g1(n)-2*g2(n)) - 3*kt^3*dz^3*g2(n)*g3(n)/C1 ) / Cf;

    TF34 = ( sqrt(3)*kt^2*dz^2*g2(n)/(g1(n)-2*g2(n)) - 2*sqrt(3)*kt*dz*g3(n)/(g1(n)-2*g2(n)) - sqrt(3)*kt^3*dz^3*g3(n)*(g1(n)-g2(n))/C1 ) / Cf;

    TF(4,1) = (-2*sqrt(3)*kt*dz*g3(n)/(g1(n)+2*g2(n))) / Cf;

    TF(4,2) = (-1 + 3*kt^2*dz^2*g3(n)^2/C1) / Cf;

    TF43 = ( sqrt(3)*kt^2*dz^2*g2(n)/(g1(n)+2*g2(n)) + 2*sqrt(3)*kt*dz*g3(n)/(g1(n)+2*g2(n)) + sqrt(3)*kt^3*dz^3*g3(n)*(g1(n)+g2(n))/C1 ) / Cf;

    TF44 = ( 2 + kt^2*dz^2*(g1(n)-g2(n))/(g1(n)+2*g2(n)) + 3*kt^3*dz^3*g2(n)*g3(n)/C1 ) / Cf;

    gp=g1(n)+2*g2(n);

    gm=g1(n)-2*g2(n);

    cons34=sqrt(3)*kt*dz^3*g3(n)/C1;

  end

  delE=V(n)-E;

  TF(3,3) = TF33 + ( dz^2*delE/gm ) / Cf;

  TF(3,4) = TF34 + ( - cons34*delE ) / Cf;

  TF(4,3) = TF43 + ( cons34*delE ) / Cf;

  TF(4,4) = TF44 + ( dz^2*delE/gp ) / Cf;


  TFtot=TF*TFtot;

  if n==N-10
    TFp=TFtot;
  end
end
```

shootvbd.m

This function looks for a minimum in the wavefunction amplitude between *Emin* and *Emax*, at intervals of *Estep*, and for a wavevector of *kt*. If a minimum is found, the minimum is determined with an accuracy *Eacc*. Optionally, the function can look for the *minnum*th minimum, disregarding the first *minnum* − 1 minima in the interval. The function works with recursive calls, refining the search grid with every step. The return values are *a*, which is the free parameter in the initial condition (*c* in the theory section), the candidate eigenenergy *E*, and a variable *error_nf*, which is set to 1 if no minimum is found and is 0 otherwise.

```
function [a,E,error_nf]=shootvbd(Emin,Emax,kt,Eacc,Estep,minnum)
%returns the approximate eigen-energy for kt, between Emin and Emax and with an accuracy Eacc

global z hbar meV V dz N g1 g2 g3

delE=Emax-Emin;

if ~exist('Estep')
```

```
    Estep=delE/10;
end

if ~exist('minnum')
    minnum=1;
end

if Estep<Eacc
    Estep=Eacc;
    fl=1;
end

E=Emin-3*Estep;
oldsignd=1;
index=0;
no_sign_change=1;
nsc=0;

while (E<=Emax) & no_sign_change
    E=E+Estep;
    index=index+1;
    tfm=tfmatrix(E,kt);
    cmin(index)=-(tfm(1,3)*tfm(1,4)+tfm(2,3)*tfm(2,4))/(tfm(1,4)^2+tfm(2,4)^2);
    diffinf(index)=(tfm(1,3)+tfm(1,4)*cmin(index))^2 + (tfm(2,3)+tfm(2,4)*cmin(index))^2;
    diffh(index)=tfm(3,3)+cmin(index)*tfm(3,4);
    diffl(index)=tfm(4,3)+cmin(index)*tfm(4,4);
    if index>1
        diff(index)=diffinf(index)-diffinf(index-1);
        newsignd=sign(diff(index));
        if index>2
        if (newsignd-oldsignd==2)
    nsc=nsc+1;
    if (nsc==minnum)
        no_sign_change=0;
    end
        end
    end
    oldsignd=newsignd;
    end
end

if no_sign_change
    a=cmin(index);
    E=E;
    error_nf=1;
    !echo No minimum found
else
    if Estep>Eacc
        [a,E,error_nf]=shootvbd(E-2*Estep,E,kt,Eacc);
    else
        a=cmin(index-1);
        E=E-Estep;
        error_nf=0;
    end
end
```

vbdos.m

This script-file cleans up some of the data generated in vbdispersion.m, and calculates the valence subband density of states as a function of in-plane hole momentum *dosk* and as a function of subband energy, relative to the valence band edge (*dosE*). The various dispersion relations are saved as .mat-files :dosk.mat, dosE.mat, rtsTE.mat, rtsTM.mat, the grid for the valence subband energy Evb.mat, the in-plane-momentum grid ktmat.mat, and the $E - k$ dispersion relation Ek.mat.

```
E_grid_step=0.1*meV;

load rtsTE
load rtsTM

% clean up E-k files

for vb=1:number_of_vb_subbands
    eval(['load kt' num2str(vb)]);
    eval(['load En' num2str(vb)]);
    indices=find(ktmat/ktstep==floor(ktmat/ktstep));
    Ek(vb,:)=subband_energy(indices);
    for cb=1:number_of_cb_subbands
        rtsTEtemp(vb,cb,:)=squeeze(rtsTE(vb,cb,indices));
        rtsTMtemp(vb,cb,:)=squeeze(rtsTM(vb,cb,indices));
    end
```

```
end

% calculate density of states vs k and DOS vs E

ktmat=[0:ktstep:ktmax];    % kt matrix
kt_num=size(ktmat,2);      % number of kt

rtsTE=rtsTEtemp;
rtsTM=rtsTMtemp;

Evb=[Emin:E_grid_step:Emax];              % energy grid for DOS vs E
Evb_num=size(Evb,2);                      % E grid size
dosE=zeros(number_of_vb_subbands,Evb_num); % init DOS vs E matrix
dEk=diff(Ek,1,2);                         % E grid step
dkt=diff(ktmat);                          % kt grid step
ktmid=ktmat(2:kt_num)-ktstep/2;           % kt grid for DOS vs kt
Ek_node_index=(Ek-Emin)/E_grid_step+1;    % index matrix to match the E grid to the kt grid

for vb=1:number_of_vb_subbands
    dosk(vb,:)=abs(ktmid.*dkt./(pi*dEk(vb,:)));                    % DOS vs kt
    for step=1:size(dEk,2)
        E1=min(Ek(vb,step),Ek(vb,step+1));
        E2=max(Ek(vb,step),Ek(vb,step+1));
        node_index_1=floor((E1-Emin)/E_grid_step+1);
        node_index_2=ceil((E2-Emin)/E_grid_step+1);
        if node_index_2<Evb_num
            for node_index=node_index_1:(node_index_2-1)
dosE(vb,node_index)=dosE(vb,node_index)+dosk(vb,step);
            end;
            node_fraction_1=(E1-Evb(node_index_1))/E_grid_step;
            dosE(vb,node_index_1)=dosE(vb,node_index_1)-node_fraction_1*dosk(vb,step);
            node_fraction_2=(Evb(node_index_2)-E2)/E_grid_step;
            dosE(vb,node_index_2-1)=dosE(vb,node_index_2-1)-node_fraction_2*dosk(vb,step);
        end
    end
end

diffdosk=diff(dosk,1,2);
dosk=[dosk(:,1) dosk(:,1:kt_num-2)+diffdosk dosk(:,kt_num-1)];

save Evb Evb
save Ek Ek
save rtsTE rtsTE
save rtsTM rtsTM
save dosk dosk
save dosE dosE
save ktmat ktmat
```

gainE.m

In this function, the gain is calculated versus the interband energy minus the bandgap energy. The gain is expressed in dimensionless numbers, the gain in cm^{-1} can be found by dividing by the module width.

The calculated gain is between valence subband vb, with population density Nb, and conduction subband cb, with population density Nc. The electron temperature is T. Optionally, a Lorentzian linewidth gam can be specified.

```
function gainE=gainE(vb,cb,Nv,Nc,T,gam)
% function gainE(vb,cb,Nv,Nc,T,gam)
% returns the gain (in %) vs (pump energy - band gap energy)
% gain in cm^-1 can be found by dividing by the module width
% uses rtsTE.mat, Ee.r, dosk.mat, ktmat.mat, Ek.mat
% constants from init.m
% vb - valence subband number
% cb - conduction subband number
% Nv - valence subband population (in units of 1e10 cm^-2)
% Nc - conduction subband population (in units of 1e10 cm^-2)
% T - temperature (K)
% gam - linewidth (default:0)

if ~exist('gam')
    gam=0;
end

init

load('Ee.r')
load rtsTE
load dosk        % DOS vs k
load Ek          % subband energy vs k
load ktmat       % kt grid
```

```
E_grid_step=0.1*meV;

Epump=[0:E_grid_step:500*meV];      % grid for pump energy in excess of band gap energy
Epump_num=size(Epump,2);            % number of nodes in Epump
gainE=zeros(Epump_num,1);           % gain vs pump E

rts=squeeze(rtsTE(vb,cb,:));
dosk_spec=squeeze(dosk(vb,:));
En=squeeze(Ek(vb,:));

gaink=q^2*hbar/(e0*c*m0^2)*rts'./(1./dosk_spec+pi*hbar^2/(0.0665*m0));  %gain vs k
gaink=gaink./(1.52*meV+En+Ee(cb,2)*meV+hbar^2*ktmat.^2/(2*0.0665*m0));
gaink=gaink*28.8*eV*m0/2;

Efv=getEfvfin(vb,Nv,T);                      % valence subband quasi-Fermi level, relative to the valence band edge
dos=1e-4*me/(pi*hbar^2);                     % density of states in cb, in cm^-2
delEf=kB*T*log(exp(1e10*Nc/(dos*kB*T))-1);   % conduction subband quasi-Fermi level, relative to the conduction subband edge
Efe=Ee(cb,2)*meV+delEf;                      % cb q-F energy, relative to the cb edge

gaink=gaink.*(fd(Efv,Ek(vb,:),T)+fd(Efe,Ee(cb,2)*meV+hbar^2*ktmat.^2/(2*0.0665*m0),T)-1);
en=Ek(vb,:)+Ee(cb,2)*meV+hbar^2*ktmat.^2/(2*0.0665*m0);   % energy difference between cb and vb vs kt

den=diff(en);              %fit gain vs k to gain vs E
for step=1:size(den,2)
  E1=min(en(step),en(step+1));
  E2=max(en(step),en(step+1));
  node_index_1=floor(E1/E_grid_step+1);
  node_index_2=ceil(E2/E_grid_step+1);
  if node_index_2<Epump_num
    for node_index=node_index_1:(node_index_2-1)
      gainE(node_index)=gainE(node_index)+gaink(step);
    end;
    node_fraction_1=(E1-Epump(node_index_1))/E_grid_step;
    gainE(node_index_1)=gainE(node_index_1)-node_fraction_1*gaink(step);
    node_fraction_2=(Epump(node_index_2)-E2)/E_grid_step;
    gainE(node_index_2-1)=gainE(node_index_2-1)-node_fraction_2*gaink(step);
  end
end

if gam~=0               % add linewidth
  factor=1/.844;   % makes up for loss due to limiting line shape width to 2*gam
  gainElw=zeros(Epump_num,1);
  abswidth=2*gam;
  delE=E_grid_step;
  Ewidth=round(abswidth/delE);
  for loop=1:size(Epump,2)
    for loop2=(loop-Ewidth):(loop+Ewidth)
      if (loop2>0) & (loop2<Epump_num)
fract=delE*(gam/(2*pi))/((Epump(loop)-Epump(loop2))^2+(gam/2)^2);
gainElw(loop2)=gainElw(loop2)+fract*gainE(loop);
      end
    end
  end
  gainE=gainElw*factor;
end
```

find_threshold_pop.m

find_threshold_pop.m recursively calculates the threshold population density for an interband transition between valence subband vb and conduction subband cb, assuming an equal population density in both subbands. The threshold gain is *threshold* (in cm^{-1}), electron temperature T. The returned population density value is in units of [1e10cm^{-2}].

```
function ftp=find_threshold_pop(vb,cb,threshold,T,popguess,popstep)
% function find_threshold_pop(vb,cb,threshold,T,popguess,popstep)
% vb = valence band number
% cb = conduction band number
% threshold = threshold gain, in % = gain in cm^-1 times module width in cm
% T = electron temperature
% popguess = initial guess for threshold population density in units of 1e10 cm^-2 (0.1)
% popstep = step size in iteration process, in 1e10 cm^-2 (5)

if ~exist('popguess')
  popguess=0.1;
end

if ~exist('popstep')
  popstep=5;
end

  tol=1e-3;
  maxgain=0;
```

```
maxsteps=250;
numsteps=1;

while (maxgain<threshold) & (numsteps<maxsteps)
    maxgain=max(gainE(vb,cb,popguess,popguess,T));
    popguess=popguess+popstep;
    numsteps=numsteps+1;
end

if (maxgain-threshold)/threshold<tol
  ftp=popguess;
else
  popstepnew=popstep/5;
  ftp=find_threshold_pop(vb,cb,threshold,T,popguess-2*popstep,popstepnew);
end
```

lifetimesfin.m

In this script file the rate equations as outlined in Chapter 4, are implemented and solved self-consistently. The e-e, h-h and LO phonon scattering rates are calculated with the aid of Paul Harrison's code. The interband generation rate is obtained from the finite-difference code. The various outputs are written to ascii data files. rhoc_vs_Txxx : conduction subband population densities rhoc_vs_Txxx : valence subband population densities eescatt_vs_Txxx : e-e scattering rates hhscatt_vs_Txxx : h-h scattering rates LO_vs_Txxx : LO phonon scattering rates tspib_vs_Txxx : interband spontaneous lifetimes and xxx denotes the photon flux in units of 10^{20}cm^{-2}.

```
clear
init
maxruns=6;
it_recalcee=5;  %value of 'it' @ which ee-scatt is recalculated

%Nc_init=load('N.r');
%Nc_init=Nc_init(:,2)*1e10;
%Nv_init=load('Nv_init.dat');
%Nv_init=Nv_init(:,2)*1e10;

Nc_init=[20 0.025 0.016 0.5]*1e10;
Nv_init=[30 2 2 2 0.01]*1e10;

for runs=1:maxruns

calculate_ee=0;   %calculate e-e scattering?
calculate_hh=0;
calculate_LO=1;   %calculate LO phonon scattering?

Msq=28.8*q*m0/2;
meV=0.001*q;
gamma=0.5*meV;
Lz=350e-8;    %effective period in cm
ibthresh=50;    %loss for depop laser in well 1 (cm^-1)
Nph=16000/(1675*meV);   %incident photon flux (cm^-2 s^-1)
sc=4;         %number of conduction band states taken into account
sv=6;         %number of valence band states taken into account
T(1)=10*runs+10;
T(2)=10*runs+10;
T(3)=10*runs+10;
T(4)=10*runs+10;
Tv(1)=10*runs+10;
Tv(2)=10*runs+10;
Tv(3)=10*runs+10;
Tv(4)=10*runs+10;
Tv(5)=10*runs+10;
Tv(6)=10*runs+10;
delta_f=4.8e11; %isb linewidth 2 meV
cc=[4 3 ; 4 2 ; 3 2];
hh=[6 2 ; 6 1 ; 5 4 ; 5 3 ; 5 2 ; 5 1 ; 4 2 ; 4 1 ; 3 5 ; 3 4 ; 3 2 ; 3 1 ; 2 1];
scatt=zeros(sc,sc);
LO=zeros(sc,sc);

!cp Nv.dat Nv_old.dat

!runsim
!cp N.r N_old.r
```

```
for it=1:10    %iteration required for self-consistent calculation

%assumed conduction band populations for first iteration
N=load('N.r');
Nc=N(:,2)*1e10;

%get fermi levels for conduction band
%load('Efe.r');
load('Ee.r');
dos=1e-4*me/(pi*hbar^2);   %density of states in cb, in cm^-2
for loop=1:sc
  delEf(loop)=kB*T(loop)*log(exp(Nc(loop)/(dos*kB*T(loop)))-1);
  Efe(loop,2)=Ee(loop,2)+delEf(loop)/meV;
end
!rm -f Efe.r
for loop=1:size(Efe,1)
  eval(['!echo ' num2str(loop) ' ' num2str(Efe(loop,2),16) ' >> Efe.r']);
end


%get fermi levels for valence subbands
Nv=load('Nv.dat');
%fd = energy diff between band edge
for loop=1:sv
  Efv(loop)=getEfvfin(loop,Nv(loop,2),Tv(loop));
      %diff between Efv and subband edge
end

%e-e scattering rates
if ((it==1) & (runs==1)) | calculate_ee
  scatt=zeros(sc,sc);
  [ree cols]=size(cc);
  for row=1:ree
    T_av=(T(cc(row,1))+T(cc(row,2)))/2;   %temp set to average of the initial subbands
    eval(['!qveescatt -T ' num2str(T_av) ' -i ' num2str(cc(row,1)) ' -f ' num2str(cc(row,2))
' -I ' num2str(N(cc(row,1),2)) ' -F ' num2str(N(cc(row,2),2)) ]);
    rr=load('temp_dir2/ccABCD.r');
    for loop=1:3
      if (rr(loop,1)==rr(loop,2)) & (rr(loop,3)==rr(loop,4))
scatt(rr(loop,1),rr(loop,4))=(scatt(rr(loop,1),rr(loop,4))+2*rr(loop,5));
      else
scatt(rr(loop,1),rr(loop,4))=(scatt(rr(loop,1),rr(loop,4))+rr(loop,5));
      end
    end
    eval(['!rm temp_dir2/*']);
    eval(['!rmdir temp_dir2']);
  end
  eval(['!rm temp_dir2/*']);
  eval(['!rmdir temp_dir2']);
  save scatt scatt;
else
  if it==1
    load scatt
    scatt_ee_init=scatt;
  end
end

%!runsimv
!cp Nv_old.dat N.r

%h-h scattering rates
if ((it==1) & (runs==1))|~| calculate_hh
  scatthh=zeros(sv,sv);
  [rhh cols]=size(hh);
  for row=1:rhh
    T_av=(Tv(hh(row,1))+Tv(hh(row,2)))/2;   %temp set to average of the initial subbands
    eval(['!qvhhscatt -T ' num2str(T_av) ' -i ' num2str(hh(row,1)) ' -f ' num2str(hh(row,2))
' -I ' num2str(Nv(hh(row,1),2)) ' -F ' num2str(Nv(hh(row,2),2)) ]);
    rr=load('temp_dir2/ccABCD.r');
    for loop=1:3
      if (rr(loop,1)==rr(loop,2)) & (rr(loop,3)==rr(loop,4))
scatthh(rr(loop,1),rr(loop,4))=(scatthh(rr(loop,1),rr(loop,4))+2*rr(loop,5));
      else
scatthh(rr(loop,1),rr(loop,4))=(scatthh(rr(loop,1),rr(loop,4))+rr(loop,5));
      end
    end
    eval(['!rm temp_dir2/*']);
    eval(['!rmdir temp_dir2']);
  end
  eval(['!rm temp_dir2/*']);
  eval(['!rmdir temp_dir2']);
  save scatthh scatthh;
else
  if it==1
    load scatthh
    scatt_hh_init=scatthh;
  end
end

!runsim
!cp N_old.r N.r

%XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX spontaneous emission lifetime %XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

load('Ee.r');
tspisb=1./zeros(sc,sc);     %lifetimes for spontaneous intersubband emission
del=zeros(sc,sc-1);         %energy diff in meV
Z=zeros(sc,sc-1);           %dipole moment
temp_vf=load('vf_e1.r');
nodes=size(temp_vf,1);       %#nodes in grid
v=zeros(nodes,sc);          %wave function amplitudes

grid=temp_vf(:,1);
v(:,1)=temp_vf(:,2);
for loop1=2:sc
  temp_vf=load(['vf_e' num2str(loop1) '.r']);
```

```
   w(:,loop1)=temp_wf(:,2);
   for loop2=1:(loop1-1)
     del(loop1,loop2)=Ee(loop1,2)-Ee(loop2,2);
     Z(loop1,loop2)=abs(trapz(grid,w(:,loop1).*grid.*w(:,loop2)));
     tspisb(loop1,loop2)=tsp_isb(del(loop1,loop2)*meV,Z(loop1,loop2));
   end
end

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXX LO scattering rates XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

load('rrp.r');
[rLO cols]=size(rrp);
if calculate_LO
  eval(['!mv rrp.r rrp_temp.r']);
  for row=1:rLO
    T_av=(T(rrp(row,1))+T(rrp(row,2)))/2;   %temp set to average of the initial subbands
    eval(['!sbp -T ' num2str(T_av)]);
    eval(['!echo ' num2str(rrp(row,1)) ' ' num2str(rrp(row,2)) ' > rrp.r']);
    eval(['!srelo -T ' num2str(T_av) ]);
    eval(['!rm rrp.r']);
  end
  eval(['!mv rrp_temp.r rrp.r']);
end
for row=1:rLO
  sbi=rrp(row,1);
  sbf=rrp(row,2);
  LOsr=load(['LOe' num2str(sbi) num2str(sbf) '.r']);
  en=LOsr(:,1)*meV;
  cst=Nc(sbi)*1e4*pi*hbar^2/me;
  LO(sbi,sbf)=trapz(en,LOsr(:,2).*fd(Efe(sbi,2)*meV,en,T(sbi)).*(1-fd(Efe(sbf,2)*meV,en-36*meV,T(sbf))))/cst;
end

LO
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXX interband generation rates and lifetimes XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

load('Eh.r');
Eval=Eh(4,2)*meV;
E_pump_exc=Eval+Ee(4,2)*meV+25*meV;%-delEexciton;
ind=round(E_pump_exc/(0.1*meV));

gain=zeros(sv,sc);
%valcon.dat : considered transitions from val --> con
load('valcon.dat');
for loop=1:size(valcon,1)
  valnr=valcon(loop,1);
  connr=valcon(loop,2);
  vnr=num2str(valnr);
  cnr=num2str(connr);
  gainvs=-gainE(valnr,connr,Nv(valnr,2),N(connr,2),T(connr));

%interband spontaneous emission rate (converted to cm^-2 s^-1)
  rrecsp(valnr,connr)=spont_em(valnr,connr,Nv(valnr,2),N(connr,2),T(connr));
  tsp(valnr,connr)=Nc(connr)/rrecsp(valnr,connr);

%calculation of pump beam absorption into each conduction subband
  gain(valnr,connr)=Nph*gainvs(ind);
end

%depop calculation
n_thresh=find_threshold_pop(1,1,ibthresh*Lz,T(1));

XXXXXXXXXXXXXXXXXXXXXXX rate equations XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

%conduction band
tc=1./zeros(sc);
for loop1=1:sc
  for loop2=1:sc
    tc(loop1,loop2)=1/(scatt(loop1,loop2)+LO(loop1,loop2)+1/tspisb(loop1,loop2));
  end
end

%valence band
tv=1./zeros(sv);
for loop1=1:sv
  for loop2=1:sv
    tv(loop1,loop2)=1/(scatthh(loop1,loop2));
  end
end

%coeff for cb
Gc=zeros(sc,1);
coeffc=zeros(sc);
for eqn=1:sc
  for coef=1:sc
    if eqn==coef
      for loop=1:sc %in cb
coeffc(eqn,eqn)=coeffc(eqn,eqn)+1/tc(eqn,loop);
      end
      for loop=1:sv %spontaneous interband emission
coeffc(eqn,eqn)=coeffc(eqn,eqn)+1/tsp(loop,eqn);
      end
    else
      coeffc(eqn,coef)=-1/tc(coef,eqn);
    end
  end;
  %spont interband emission
  for loop=1:sv
    Gc(eqn)=Gc(eqn)+gain(loop,eqn);
  end
end
rhoc=coeffc^-1*Gc;
```

```
%coeff for vb
Gv=zeros(sv,1);
coeffv=zeros(sv);
for eqn=1:sv
  for coef=1:sv
    if eqn==coef
      for loop=1:sv %in vb
coeffv(eqn,eqn)=coeffv(eqn,eqn)+1/tv(eqn,loop);
      end
      for loop=1:sc %spontaneous interband emission
coeffv(eqn,eqn)=coeffv(eqn,eqn)+1/tsp(eqn,loop);
      end
    else
      coeffv(eqn,coef)=-1/tv(coef,eqn);
    end
  end;
  %spont interband emission
  for loop=1:sc
    Gv(eqn)=Gv(eqn)+gain(eqn,loop);
  end
end
rhov=coeffv^-1*Gv;

if rhoc(1)>n_thresh*1e10
  rhoc(1)=n_thresh*1e10;
end;
if rhov(1)>n_thresh*1e10
  rhov(1)=n_thresh*1e10;
end;


!rm N.r
!rm Nv.dat
for loop=1:sc %adjust ee-rates
  eval(['!echo ' num2str(loop) ' ' num2str(rhoc(loop)/1e10) ' >> N.r']);
  scatt(loop,:)=scatt_ee_init(loop,:)*Nc(loop)/Nc_init(loop);
end
for loop=1:sv %adjust hh-rates
  eval(['!echo ' num2str(loop) ' ' num2str(rhov(loop)/1e10) ' >> Nv.dat']);
  scatthh(loop,:)=scatt_hh_init(loop,:)*(Nv(loop,2)*1e10)/Nv_init(loop);
end


calculate_LO=0;
calculate_ee=0;
calculate_hh=0;
if it==it_recalcee
  calculate_ee=1;
  calculate_hh=1;
end

end %iteration


%intersubband gain in cm^-1
isbgain43=(rhoc(4)-rhoc(3))*q^2*del(4,3)*meV*(Z(4,3)*1e2)^2/(pi*n*e0*c*hbar^2*delta_f*Lz);
isbgain42=(rhoc(4)-rhoc(2))*q^2*del(4,2)*meV*(Z(4,2)*1e2)^2/(pi*n*e0*c*hbar^2*delta_f*Lz);


eval(['!echo ' num2str(T) '  ' num2str(transpose(rhoc)/1e10) ' >> rhoc_vs_T' num2str(Nph/1e20)]);

eval(['!echo ' num2str(T) '  ' num2str(transpose(rhov)/1e10) ' >> rhov_vs_T' num2str(Nph/1e20)]);

eval(['!echo ' num2str(T) ' 3 2 ' num2str(1/scatt(3,2),16) ' >> eescatt_vs_T' num2str(Nph/1e20)]);

eval(['!echo ' num2str(T) ' 4 2 ' num2str(1/scatt(4,2),16) ' >> eescatt_vs_T' num2str(Nph/1e20)]);

eval(['!echo ' num2str(T) ' 4 3 ' num2str(1/scatt(4,3),16) ' >> eescatt_vs_T' num2str(Nph/1e20)]);

for loop=1:size(hh,2)
  v1=hh(loop,1);
  v2=hh(loop,2);
  eval(['!echo ' num2str(T) ' ' num2str(v1) ' ' num2str(v2) ' ' num2str(1/scatthh(v1,v2),16) ' >> hhscatt_vs_T' num2str(Nph/1e20)]);
end


for row=1:rLO
  sbi=rrp(row,1);
  sbf=rrp(row,2);
  eval(['!echo ' num2str(T) '  ' num2str(sbi) ' ' num2str(sbf) ' ' num2str(1/LO(sbi,sbf),16) ' >> LO_vs_T' num2str(Nph/1e20)]);
end

for loopv=1:sv
  for loopc=1:sc
    vnr=num2str(loopv);
    cnr=num2str(loopc);
    eval(['!echo ' num2str(T) ' ' vnr ' ' cnr ' ' num2str(tsp(loopv,loopc)) ' >> tspib_vs_T' num2str(Nph/1e20)]);
  end
end

eval(['!echo ' num2str(T) ' 4 2 ' num2str(isbgain42,16) ' >> isbgain_vs_T' num2str(Nph/1e20)]);
eval(['!echo ' num2str(T) ' 4 3 ' num2str(isbgain43,16) ' >> isbgain_vs_T' num2str(Nph/1e20)]);

eval(['!echo ' num2str(Nph) ' ' num2str(isbgain43,16) ' >> isbgain43_vs_G']);
eval(['!echo ' num2str(Nph) ' ' num2str(isbgain42,16) ' >> isbgain42_vs_G']);


end %runs
```

Bibliography

[1] M. Tacke. New developments and applications of tunable IR lead salt lasers. *Infrared Phys. Technol.*, 36:447–463, 1995.

[2] S.E. Rosenbaum, B.K. Kormanyos, L.M. Jelloin, M. Matloubian, A.S. Brown, E. Larson, L.D. Nguyen, M.A. Thompson, L.P.B. Katehi, and G.M. Rebeiz. 155- and 213-Ghz AlInAs/GaInAs/InP HEMT MMIC oscillators. *IEEE Trans. Microwave Theory Tech.*, 43:927–932, 1995.

[3] L. Esaki and R. Tsu. Superlattice and negative differential conductivity in semiconductors. *IBM J.Res.Dev.*, 14:61, 1970.

[4] R.F. Kazarinov and R.A. Suris. Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice. *Soviet Physics: Semiconductors*, 5:707–709, 1971.

[5] R. Colombelli, F. Capasso, C. Gmachl, A.L. Hutchinson, D.L. Sivco, A. Tredicucci, M.C. Wanke, A.M. Sergent, and A.Y. Cho. Far-infrared surface-plasmon quantum-cascade lasers at 21.5 μm and 24 μm wavelengths. *Appl. Phys. Lett.*, 78:2620–2622, 2001.

[6] J. Faist, F. Capasso, C. Sirtori, D. Sivco, J.N. Baillargeon, A.L. Hutchinson, and A.Y. Cho. High power mid-infrared ($\lambda \approx 5\mu$m) quantum cascade lasers operating above room temperature". *Appl.Phys.Lett.*, 68:3680–3682, 1996.

[7] J. Faist, C. Sirtori, F. Capasso, D. Sivco, J.N. Baillargeon, A.L. Hutchinson, and A.Y. Cho. High power long wavelength ($\lambda \approx 11.5\mu$m) quantum cascade lasers operating above room temperature. *Photon. Technol. Lett.*, 10:1100–1102, 1998.

[8] O. Gauthier-Lafaye, P. Boucaud, F.H. Julien, S. Sauvage, S. Cabaret, J.M. Lourtioz, V. Thierry-Mieg, and R. Planel. Long wavelength ($\approx 15.5\mu$m) unipolar semiconductor laser in GaAs quantum wells. *Appl. Phys. Lett.*, 71:3619–3621, 1997.

[9] S. Sauvage, Z. Moussa, P. Boucaud, F.H. Julien, V. Berger, and J. Nagle. Room temperature infrared intersubband photoluminescence in GaAs quantum wells. *Appl. Phys. Lett.*, 70:1345–1347, 1997.

[10] S. Sauvage, P. Boucaud, F.H. Julien, O. Gauthier-Lafaye, V. Berger, and J. Nagle. Intersubband photoluminescence of GaAs quantum wells under selective interband excitation. *Appl. Phys. Lett.*, 71:1183–1185, 1997.

[11] I. Lyubomirsky, Q. Hu, and R. Melloch. Measurement of far-infrared intersubband spontaneous emission from optically pumped quantum wells. *Appl. Phys. Lett.*, 73:3043–3045, 1998.

[12] F.H. Julien, A. Sa'ar, J. Wang, and J.P. Leburton. Optically pumped intersubband emission in quantum wells. *Appl. Phys. Lett.*, 71:3619–3621, 1997.

[13] T.J. Green and W. Xu. Population inversion in an optically pumped single quantum well. *J. Appl. Phys.*, 88:3166–3169, 2000.

[14] A. Kastalsky. Infrared interband laser induced in a multiple-quantum-well interband laser. *IEEE J. Quantum Elec.*, 4:1112–1115, 1993.

[15] I. Vurgaftman, J.R. Meyer, F.H. Julien, and L.R. Ram-Mohan. Design and simulation of low-threshold antimonide intersubband lasers. *Appl. Phys. Lett.*, 73:711–713, 1998.

[16] H.Q. Le, W.D. Goodhue, P.A. Maki, and S. Di Cecca. Diode-laser pumped InGaAs/GaAs/AlGaAs heterostructure lasers with low internal loss and 4w average power". *Appl.Phys.Lett.*, 63:1465–1467, 1993.

[17] S.L. Chuang. *Physics of optoelectronic devices*. John Wiley and Sons, New York, 1995.

[18] J.M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Phys. Rev.*, 97:869–883, 1955.

[19] J.M. Luttinger. Quantum theory of cyclotron resonance in semiconductors : General theory. *Phys. Rev.*, 102:1030–1041, 1956.

[20] R. Eppenga, M.F.H. Schuurmans, and S. Colak. New k.p theory for GaAs/AlGaAs-type quantum wells. *Phys.Rev.B*, 36:1554–1564, 1987.

[21] D.A. Broido and L.J. Sham. Effective masses of holes at GaAs-AlGaAs heterojunctions. *Phys. Rev. B*, 31:888–892, 1985.

[22] R.A. Stradling. The application of high magnetic fields in semiconductor physics. *Proc.Confer.*, page 257, 1974.

[23] P. Lawaets. *Phys.Rev. B*, 4:3460, 1971.

[24] I.M. Tsidilkovsky. *Band structure of semiconductors*. Pergamon Press, Oxford, UK, 1982.

[25] J.S. Blakemore. Semiconducting and other major properties of gallium arsenide. *J. Appl. Phys.*, 53:R123–R179, 1982.

[26] S. Adachi. GaAs, AlAs and $Al_xGa_{1-x}As$: material parameters for use in research and device applications. *J. Appl. Phys.*, 58:R1–R29, 1985.

[27] L.C. Andreani, A. Pasquarello, and F. Bassani. Hole subbands in strained GaAs-Al$_x$Ga$_{1-x}$As quantum wells: exact solution of the effective-mass equation. *Phys. Rev.*, 102:1030–1041, 1956.

[28] L.A. Coldren and S.W. Corzine. *Diode Lasers and Photonic Integrated Circuits*. John Wiley and Sons, New York, 1995.

[29] G. Mahan. *Many-particle Physics*. Plenum, New York, 1990.

[30] J.H. Smet. *Intrawell and interwell intersubband transitions in single and multiple quantum well heterostructures*. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1995.

[31] C.J. Hearn. *The physics of nonlinear transport in semiconductors, edited by D.K. Ferry, J.R. Barker and C. Jacoboni*, pages 153–166. Plenum, New York, 1980.

[32] P.F. Maldague. Many-body corrections to the polarizability of the two-dimensional electron gas. *Surf.Sci.*, 73:296, 1978.

[33] T. Elsaesser, J. Shah, L. Rota, and P. Lugli. Initial thermalization of photoexcited carriers in GaAs studied by femtosecond luminescence spectroscopy. *Phys.Rev.Lett.*, 66:1757–1760, 1991.

[34] W. Knox. Optical studies of femtosecond carrier thermalization. *Hot carriers in semiconductor nanostructures, edited by J. Shah*, pages 313–344, 1992.

[35] J. Shah, A. Pinczuk, A.C. Gossard, and W. Wiegmann. Energy loss rates for hot electrons and holes in GaAs quantum wells. *Phys.Rev. Lett.*, 54:2045–2048, 1985.

[36] J. Shah. Hot carriers in 2-D polar semiconductors. *IEEE J.Quantum Elec.*, 22:1728–1743, 1986.

[37] J. Shah. Hot electrons and phonons under high intensity photoexcitation of semiconductors. *Solid-State Electronics*, 21:43–50, 1978.

[38] T.B. Norris, N. Vodjdani, B. Vinter, E. Costard, and E. Böckenhoff. Resonant tunneling between heavy-hole and light-hole states in coupled quantum wells. *Phys.Rev.B*, 43:1867–1870, 1991.

[39] P. Harrison. *Quantum wells, wires and dots*. John Wiley and Sons, Ltd, Chichester, UK, 1999.

[40] B. Xu. *Development of Intersubband Terahertz Lasers Using Multiple Quantum Well Structures*. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Cambridge, MA, 1998.

[41] Edward D. Palik. *Handbook of Optical Constants of Solids*, pages 429–443. Academic Press, Orlando, FL, 1985.