

## *The implementation of a drift diffusion based thermal model*

### **5.1 Introduction**

The optimisation of optoelectronic devices can be costly, requiring multiple MBE growths and hours of processing. A way to avoid this expensive process is to use truly predictive device simulation tools firstly to understand the physical processes within the device and then to optimise the device design. One of the most important aspects of a device simulator is its ability to accurately predict the device temperature. Almost all material parameters depend on device temperature, mobility, thermal conductivity, gain and band gap energy to name but a few. It is therefore essential to accurately include thermal effects if realistic models are to be created.

In this chapter, our isothermal device simulator is extended to include thermal effects. The thermal device equations are derived, discretised and then implemented. A material database is designed and implemented to enable material parameters to be quickly calculated as a function of temperature. The interaction of the electrical and thermal meshes is investigated. Finally, various numerical stabilisation techniques are evaluated to ensure the efficient and stable solution of the electro-thermal model.

## 5.1 The lattice heat equation

The classical description of heat flow in solids was first published by Fourier in his paper “*On the Propagation of Heat in Solid Bodies*” in 1807. Fourier's heat equation for solids is usually written as

$$\mathbf{S}_L(\mathbf{r}, t) = -\kappa_L \nabla T_L(\mathbf{r}, t) \quad , \quad (5.1)$$

where  $S_L$  is the heat flux density,  $k_L$  is the lattice thermal conductivity and  $T_L$  is the lattice temperature. This equation states that the heat flux density is proportional to the gradient of the temperature multiplied by the thermal conductivity.

To derive the more general heat diffusion equation, the conservation of heat is considered. If a volume element has a heat flux  $S_L$  flowing through its walls, the integral of this energy flux out of the volume element plus the total change in energy stored within the unit volume is equal to the total heat generation within the volume.

This may be written as [1]

$$\oint_A \mathbf{S}_L(\mathbf{r}, t) \cdot d\mathbf{A} = -\frac{\partial}{\partial t} \int_V \rho_L c_L T_L dV + \int_V H dV \quad , \quad (5.2)$$

where  $t$  is time,  $\rho_L$  is the density of the material,  $c_L$  is the specific heat at a constant pressure and  $H(T)$  is the rate at which heat is generated per unit volume. Then using Gauss's law

$$\oint_A \mathbf{S}_L d\mathbf{A} = \int \nabla \cdot \mathbf{S}_L dV \quad (5.3)$$

with 5.1, 5.2 and differentiating with respect to volume, the heat diffusion equation is obtained as

$$\rho_L c_L \frac{\partial T(\mathbf{r})}{\partial t} = \nabla \cdot \kappa_L \nabla T(\mathbf{r}) + H(T, \mathbf{r}) . \quad (5.4)$$

The term on the left hand side of the equation describes heat storage. The two terms on the right hand side describe the heat flow into and out of a unit volume and heat generation or loss within that unit volume. As long as the device is at steady state, the time dependence can be neglected, i.e.

$$\frac{\partial T_L}{\partial t} = 0 . \quad (5.5)$$

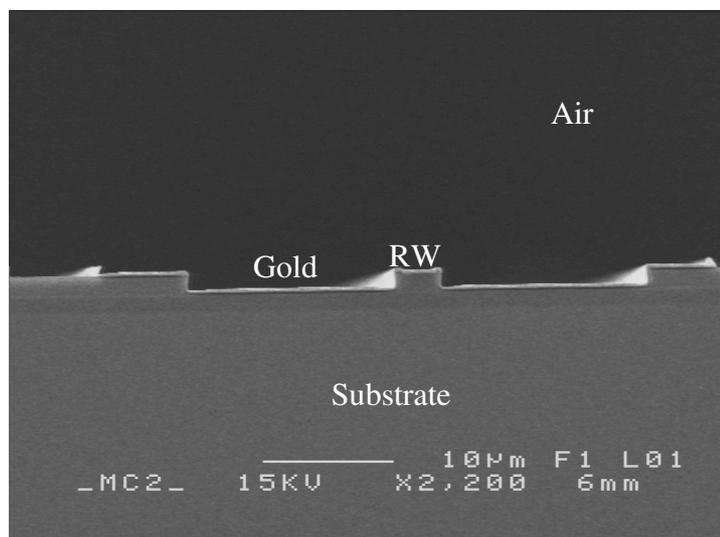
Thus, equation 5.4 becomes,

$$0 = \nabla \cdot \kappa_L \nabla T_L + H , \quad (5.6)$$

which is the steady state lattice heat equation.

## 5.2 Thermal boundary conditions

A key aspect of solving the heat equation is the choice of the correct boundary conditions. A typical  $1.3\mu\text{m}$  ridge waveguide (RW) laser diode is shown below in figure 5.1. Thermally, there are three types of boundary conditions which can be used to model such a problem - *Dirichlet*, *Neumann* and *mixed conditions*. These boundary conditions are described in the following section.



*Figure 5.1: An SEM image of a  $1.3\mu\text{m}$  laser diode. (Image courtesy of Chalmers University of Technology, Sweden.)*

### 5.2.1 Neumann boundary conditions

A Neumann boundary condition sets the value of the derivative at the material interface. In this case, it is used to force the heat flux across an interface to zero. One place where such a boundary condition might be applied is the interface between the air and the gold in figure 5.1. To achieve zero heat flux across the interface, the

gradient of the temperature must be 0 normal to the interface, i.e.

$$\mathbf{S} \cdot \hat{\mathbf{n}} = k(\hat{\mathbf{n}} \cdot \nabla T) = 0 \quad , \quad (5.7)$$

where  $\hat{\mathbf{n}}$  is a unit vector normal to the surface. For this to be true, the temperature on either side of the interface must be forced to be the same.

### 5.2.2 Dirichlet boundary conditions

The Dirichlet boundary condition sets the temperature at the edge of the simulation window to a constant value. This would be used to model the interface of the substrate with a stabilised Peltier cooler or heat sink with a very low thermal resistance.

### 5.2.3 Mixed boundary conditions

Mathematically, mixed boundary conditions describe a boundary where the boundary conditions change along the edge of the solution space. These boundary conditions can be used to model a heat sink without the computational overhead of extending the mesh over the entire heat sink.

Figure 5.2 depicts the device heat sink interface, where  $T_l$  is the last mesh point within the device,  $T_3$  the external temperature of the heat sink (commonly set to 300K) and  $T_2$  a temperature point on the boundary between the heat sink and the device. Temperature  $T_2$  must be calculated.

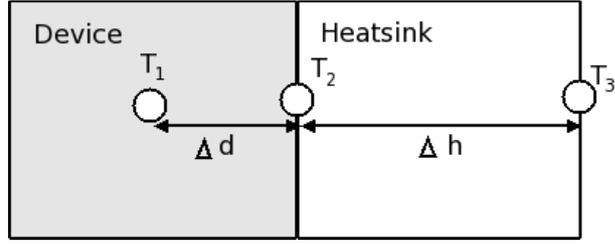


Figure 5.2: Device heat sink contact.

By forcing conservation of heat flux normal to the interface,

$$\vec{q}_{heat\ sink} \cdot \hat{n} = \vec{q}_{device} \cdot \hat{n} \quad , \quad (5.8)$$

Fourier's equation for the heat sink-device interface may be written as

$$-k_{device} \frac{\partial T_{device}}{\partial x} = -k_{heat\ sink} \frac{\partial T_{heat\ sink}}{\partial x} \quad . \quad (5.9)$$

Writing in terms of  $T_1$ ,  $T_2$  and  $T_3$ , this gives

$$-k_d \frac{T_2 - T_1}{\Delta d} = -k_h \frac{T_2 - T_3}{\Delta h} \quad , \quad (5.10)$$

where  $\Delta h$  is the length of the heat sink,  $\Delta d$  is the distance between the two mesh points closest to the heat sink,  $k_d$  is the thermal conductivity of the device and  $k_h$  is the effective thermal conductivity of the heat sink. The equation may be rearranged to give the temperature  $T_2$

$$T_2 = \frac{T_3 k_h \Delta d + T_1 \Delta h k_d}{k_d \Delta h + k_h \Delta d} \quad . \quad (5.11)$$

$k_h$  is calculated from the heat sink thermal resistance using

$$k_h = \frac{\Delta h}{A_{th} R_{th}} \quad , \quad (5.12)$$

where  $A_{th}$  is the contact area and  $R_{th}$  (Kelvin/Watt) is the thermal resistance of the heat sink. The effective conductivity of the heat sink  $k_h$  is much higher than the thermal

conductivity of the material from which the heat sink is made. This is because the Neumann boundary condition assumes a 1D heat flux. In reality, however, the heat sink is a 3D structure with 3D heat flow. Thus, heat escapes from all sides of the heat sink not just one as the equation suggests.

Although the heat sink can be modelled in this way, there is significant lateral heat flow within the first few microns of the heat sink. This lateral heat flow spreads the heat across the top of the device and alters the thermal profile within the device itself. The mesh was therefore extended over the first few microns ( $\sim 2\text{-}3\mu\text{m}$ ) of the heat sink in order to include this effect.

### 5.3 Discretisation of the thermal problem

The lattice heat equation (5.6) must be discretised before it can be solved

$$0 = \nabla \cdot \kappa \nabla T + H \quad . \quad (5.6)$$

In 1D the discrete form of the derivative of the heat flux (in equation 5.6) is

$$\frac{\partial}{\partial x} \cdot \left( \kappa \frac{\partial T}{\partial x} \right) = \frac{-k_1 \frac{(T_1 - T_0)}{\Delta x_1} + k_2 \frac{(T_2 - T_1)}{\Delta x_2}}{\frac{\Delta x_1}{2} + \frac{\Delta x_2}{2}} + O \quad (5.13)$$

where,  $T_{0,1,2}$  and  $k_{0,1,2}$  are the temperatures and thermal conductivities associated with calculation cell 0, 1 and 2, respectively (see figure 5.3). Equation 5.13 can be rewritten as

$$\frac{\partial}{\partial x} \cdot \kappa \frac{\partial}{\partial x} = \frac{2 \Delta x_2 k_1 T_0 - 2(k_1 \Delta x_2 + k_2 \Delta x_1) T_1 + 2k_2 \Delta x_1 T_2}{(\Delta x_1 + \Delta x_2) \Delta x_1 \Delta x_2} + O \quad , \quad (5.14)$$

where  $O$  is the remainder due to the imperfect approximation of 5.14. The same procedure can be repeated in the  $y$  direction and substituted into equation 5.6, forming the 2D discretised heat equation for a mesh with a variable cell size

$$\frac{2 \Delta x_2 k_1 T_0 - 2(k_1 \Delta x_2 + k_2 \Delta x_1) T_1 + 2k_2 \Delta x_1 T_2}{(\Delta x_1 + \Delta x_2) \Delta x_1 \Delta x_2} + \frac{2 \Delta y_2 k_3 T_0 - 2(k_3 \Delta y_2 + k_4 \Delta y_1) T_1 + 2k_4 \Delta y_1 T_2}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} + H = 0 \quad . \quad (5.15)$$

It is worth noting that the derivative is often incorrectly evaluated, by removing the thermal conductivity term  $k$ , from inside the divergence, i.e.

$$\nabla \cdot \kappa \nabla T \neq \kappa \nabla \cdot \nabla T \quad . \quad (5.16)$$

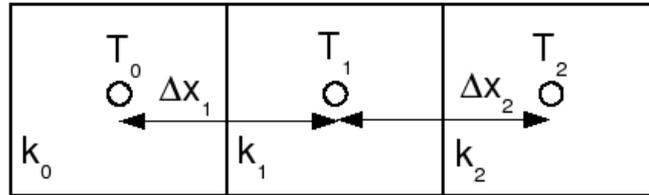


Figure 5.3: An example of a calculation mesh.

#### 5.4 Coupled iterative solution of the electrical and thermal problem

The solution method chosen was a coupled iterative method, whereby the heat equation is solved outside of the main Newton solver for the electrical equations. The electrical and thermal solvers are run after one another consecutively until

convergence is achieved. This was found to be a stable and robust method for the solution of the electro-thermal problem.

#### **5.4.1 The 1D problem**

In order to solve the 1D thermal problem, the electrical solver is first called and run to convergence. The heat sources and thermally dependent parameters are then updated. The thermal solver is then run until convergence and passes the thermal profile back to the electrical solver, which updates the thermally dependant electrical parameters. The process is then repeated until a self consistent solution is found. A flow diagram of the solver is shown in figure 5.4. This usually takes up to ten iterations back and forth between the electrical and thermal solvers to obtain an extremely well converged solution. In the latter stages of solution, the individual solvers converge very quickly because the initial guess is close to the correct answer. Thus, the whole process only takes two or three times as long as an isothermal simulation (i.e. where the temperature in the device is assumed to be uniform and constant).

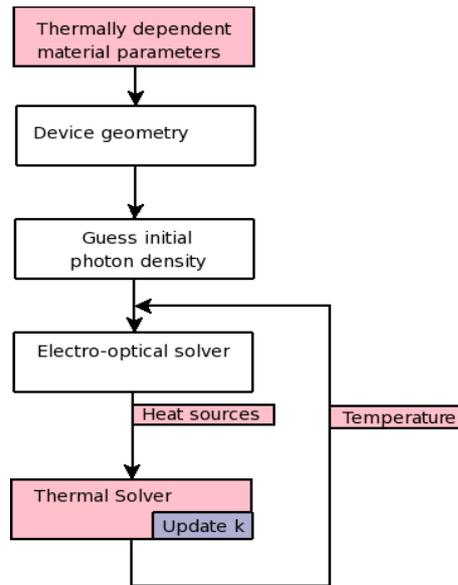


Figure 5.4: Flow diagram of the coupled solution method.

#### 5.4.2 The 2D problem

Once the 1D simulator has been run to convergence the result is used as an initial guess for the 2D simulator. The electrical solver solves the 2D electrical problem in one large sparse matrix. This takes a considerable time to build and solve. The problem also takes far more iterations to solve than the 1D problem. Thus, it was found computationally wasteful to run the electrical solver to convergence before updating solving the thermal problem. It is more efficient to solve the electrical solver for a few iterations (~6) and then solve thermal problem fully before returning to the electrical problem. The thermal solver only takes one or two iterations to fully converge. By using this approach, the slow electrical solver does not waste time solving for an inaccurate thermal profile.

#### 5.5 Solving the 2D thermal problem

### 5.5.1 Alternating Direction Implicit methods (ADI)

The Alternating Direction Implicit method was evaluated as a solution method for the thermal problem. In ADI, a 2D problem is solved by solving 1D strips of the problem. The solver scans the problem space solving 1D strips in one direction (e.g. vertical), followed by 1D strips in the other direction (e.g. horizontal). This process is repeated until convergence is reached. This method is commonly used in the solution of optical and electrical problems. The advantage of this method is that the entire matrix must not be solved at once, but the disadvantage is that it is slow due to its iterative nature. For the solution of the heat equation, ADI was found to be very slow – in part since it is hard to give a good initial guess for the thermal problem.

### 5.5.2 Gaussian elimination

By writing equation 5.15 for every mesh point, a set of linear equations is formed. Although it is possible to solve linear equations via Gaussian elimination, this would remove any possibility of extending the thermal model later to include non linear terms (e.g. non-equilibrium heating mechanisms in the QW).

### 5.5.3 Newton's method

One of the most powerful methods used for the solution of both linear and non-linear equations is Newton's method. To implement this method, an error function is defined,

$$\frac{\partial}{\partial x} k \frac{\partial}{\partial x} T + H = W(T) \quad . \quad (5.17)$$

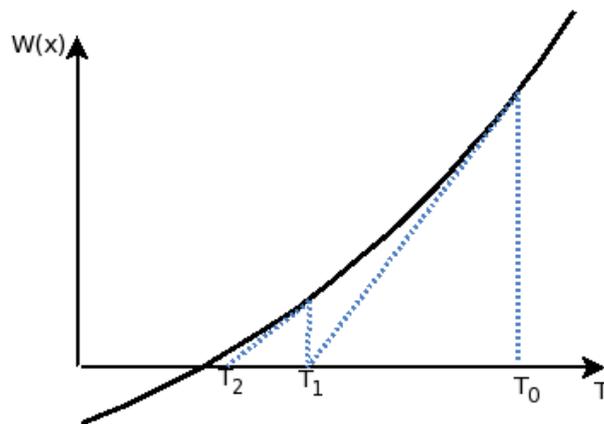
The equation has been solved when  $W(T)=0$ . The derivative of 5.17 with respect to  $T$  is given as  $W'(T)$ . From figure 5.5, it can be seen that

$$W'(T_0) = \frac{W(T_0)}{T_0 - T_1} \quad (5.18)$$

This can be rearranged to give

$$T_1 = T_0 - \frac{W(T_0)}{W'(T_0)} \quad (5.19)$$

where  $T_1$  is a guess closer to the solution of the problem. This process is repeated until a sufficiently small error is obtained (see figure 5.5).



*Figure 5.5: An example of error  $W(T)$  plotted against  $T$  [1].*

Newton's method can be extended to systems of equations, by rewriting equation 5.18 in matrix form

$$\left[ \frac{\partial W(T)}{\partial T} \right] [\Delta T] = [W] \quad (5.20)$$

This matrix equation is solved to obtain the matrix  $[\Delta T]$ , which is then subtracted

from the matrix  $[T]$ . This process is repeated until convergence is achieved.

### 5.5.4 Solving the 1D heat equation using Newton's method

To solve the heat equation in 1D using Newton's method, the discretised form of equation 5.6) must be written as an error function

$$\frac{2 \Delta y_2 k_3 T_{j-1} - 2(k_3 \Delta y_2 + y_2 \Delta y_1) T_1 + 2k_4 \Delta y_1 T_{j+1}}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} + H = W_j(T) \quad .$$

(5.21)

The derivative with respect to temperature for each mesh point must then be taken,

$$d_j = \frac{\partial W_j(T)}{\partial T_{j-1}} = \frac{2 \Delta y_2 k_3}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} \quad (5.22)$$

$$a_j = \frac{\partial W_j(T)}{\partial T_j} = \frac{-2(k_3 \Delta y_2 + y_2 \Delta y_1)}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} \quad (5.23)$$

$$b_j = \frac{\partial W_j(T)}{\partial T_{j+1}} = \frac{2k_4 \Delta y_1}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} \quad (5.24)$$

The set of equations 5.21-5.24 can be written for every mesh point  $0 \leq j < M$

$$\begin{bmatrix} a_0 & b_0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 \\ d_1 & a_1 & b_1 & \dots & \dots & \dots & \dots & 0 & 0 & 0 \\ & d_2 & a_2 & b_2 & \dots & \dots & \dots & \dots & 0 & 0 \\ & \dots & d_3 & a_3 & b_3 & \dots & \dots & \dots & \dots & 0 \\ & \dots & \dots & \ddots & \ddots & \ddots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & d_{M-4} & a_{M-4} & b_{M-4} & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & d_{M-3} & a_{M-3} & b_{M-3} & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & d_{M-2} & a_{M-2} & b_{M-2} & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & \dots & d & a_{M-1} & b_{M-1} & \dots \end{bmatrix} \begin{bmatrix} \Delta T_0 \\ \Delta T_1 \\ \Delta T_2 \\ \Delta T_3 \\ \dots \\ \Delta T_{M-4} \\ \Delta T_{M-3} \\ \Delta T_{M-2} \\ \Delta T_{M-1} \end{bmatrix} = \begin{bmatrix} W_0 \\ W_1 \\ W_2 \\ W_3 \\ \dots \\ W_{M-4} \\ W_{M-3} \\ W_{M-2} \\ W_{M-1} \end{bmatrix} \quad (5.25)$$

For a Dirichlet boundary condition, at mesh point  $0$  or  $M-1$ , the following derivatives are used

$$d_{-1}=0 \quad (5.26)$$

$$b_M=0 \quad (5.27)$$

respectively. For a Neumann boundary condition, the temperature in the cell just outside the mesh  $d_{-1}$  or  $d_M$  is the same as the temperature in the cell just inside the mesh. For fast convergence, the derivative in cell  $d_0$  or  $d_{M-1}$  must be modified to reflect this. Thus, for Neumann conditions at  $j=0$

$$a_0 = \frac{2 \Delta y_1 k_3}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} - \frac{2(k_3 \Delta y_2 + y_2 \Delta y_1)}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} \quad (5.28)$$

or in cell  $j=M-1$

$$b_{M-1} = \frac{2k_4 \Delta y_1}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} - \frac{2(k_3 \Delta y_2 + y_2 \Delta y_1)}{(\Delta y_1 + \Delta y_2) \Delta y_1 \Delta y_2} . \quad (5.29)$$

Equation 5.25 is a tridiagonal sparse matrix, which is solved for the update  $[\Delta T]$ . This  $[\Delta T]$  is then added to the initial guess for  $[T]$  and the process is repeated until convergence is achieved. Convergence is defined as when the summation of the moduli of the error functions is smaller than a predefined error  $E$ .

$$\sum_{n=1}^N |W_n| < E . \quad (5.30)$$

For a typical simulation,  $E$  was set to  $1 \times 10^{-6}$ .

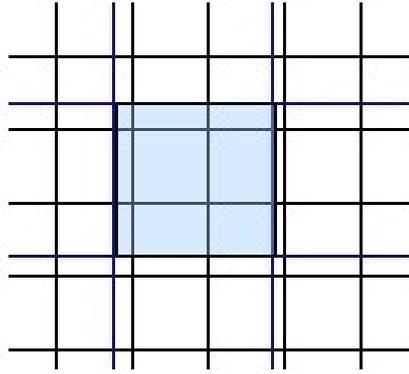
### 5.5.5 Solving the 2D heat equation



The use of independent electrical and thermal meshes allows both meshes to be independently fine tuned. The temperature varies relatively slowly as a function of position. Thus, a linear interpolation algorithm can be used to interpolate the temperature profile between the thermal and electrical meshes. However, due to the complex epitaxial structure of a laser diode, the heat sources do not vary smoothly. This is especially true at heterojunction interfaces and around the QW. Therefore, to interpolate the heat sources between the meshes, a different interpolation algorithm must be used. One method is to use the weighted mean of the heat sources. Within a thermal mesh square, the weighted mean may be written as

$$H_{i,j} = \sum_{jj=0}^{M_e} \sum_{ii=0}^{N_e} \frac{f(i,j,ii,jj) \cdot H_{ii,jj}}{A_{i,j}}, \quad (5.32)$$

where  $M_e$  and  $N_e$  are the number of points on the  $j$  and  $i$  axes of the electrical mesh,  $f$  is a function which is equal to the area of the electrical mesh square which lies within the thermal mesh square,  $H_{ii,jj}$  is the heat source at point  $ii,jj$  on the electrical mesh and  $A_{i,j}$  is the area of the thermal mesh square at  $i,j$ . To evaluate  $f$  for each thermal mesh point, the surrounding area in the electrical mesh must be searched, and for every electrical mesh point, it must be determined how much of the cell lies within the thermal cell. This process is depicted in figure 5.6.



*Figure 5.6: An example of an overlapping electrical and thermal mesh. The black mesh represents the fine electrical mesh, while the blue mesh represents the thermal mesh. The blue square represents the overlap of a thermal cell with the electrical cells.*

Due to the amount of searching and calculation required to evaluate  $f$ , this algorithm was found to be very slow. Another problem found with the interpolation method is that the electrical problem is very strongly dependent on temperature. For example, carrier density, mobility, gain and stimulated recombination are temperature dependent as are many other parameters. Thus, significant errors can be introduced into the model by linearly interpolating the temperature.

### **5.6.2 Using the same electrical and thermal mesh**

If the thermal mesh points are made to coincide with the electrical mesh points, no interpolation is required. However, the thermal problem often needs to be solved in regions extending beyond the electrically active regions, which include the substrate, contacts, heat sink and etched trenches. Thus, the thermal mesh was extended past

the bounds of the electrical mesh giving a hybrid thermal mesh. This approach was found to be the most efficient and most stable.

### 5.7 The materials database

During the initialisation of the original *isothermal* laser simulator, the material parameters were read directly from a materials database for each mesh point. After this, the constants (e.g. mobility) were not changed. Thus, there was no need to maintain a copy of the complex polynomials describing the material parameters in memory. If the simulator were to be made temperature dependent, material constants such as mobility and thermal conductivity would have to be updated as the internal temperature changed during the course of the simulation. Rereading the data from the database on disk at every mesh point was a very slow process. Consequently, a system had to be developed for storing the material parameters in memory from which the materials constants could be recalculated quickly and efficiently. The class structure of this storage system is shown in figure 5.7. On initialisation of the laser simulator, the material database is read for each material layer. A class structure is then dynamically generated and populated using the material constants. A *material* class is associated with each epitaxial layer (figure 5.7). The *material* class contains an array of class *matparam*, one for each material parameter (e.g. thermal conductivity, mobility, etc.). Each *matparam* class contains the parameters to describe the material parameter as a function of material composition, temperature and doping density. When a material parameter is required for a mesh point, the function *get\_value* is called and returns the required material parameter.

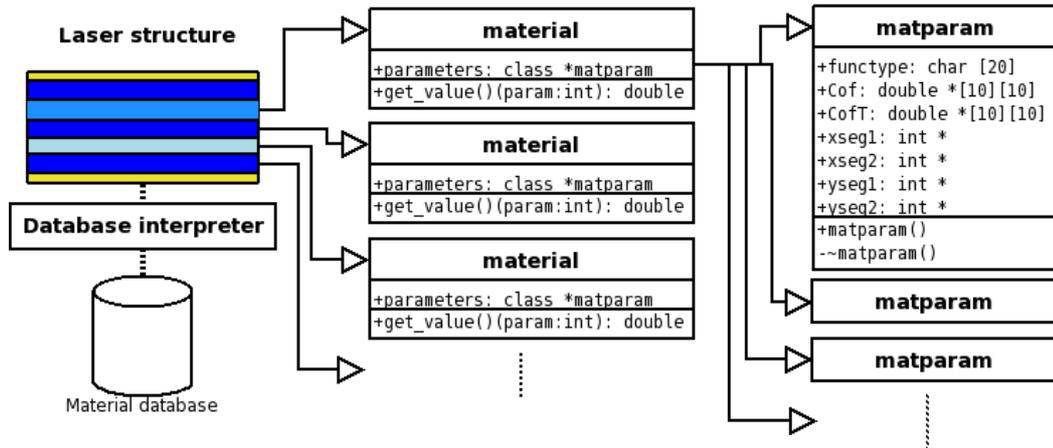


Figure 5.7: A simplified diagram describing the structure and inheritance of the materials database classes.

The temperature dependences of the material models had to be found and the materials database routines had to be rewritten to handle the data. The file format used to store the materials data base was kept broadly compatible with CONAN, a first generation laser simulation tool [2] written by the University of Nottingham and the Universidad Politécnica de Madrid. Another feature of the new database was the ability of the database files to accept comment lines. Thus, a list of references as to where in the literature the parameters came from could be kept with the material parameters.

## 5.8 Material models

### 5.8.3 Bulk carrier densities

Boltzmann statistics are used to model the 3D carrier densities. The temperature dependent carrier density for electrons is given as

$$n = N_c \exp\left(\frac{F_n - E_c}{kT_e}\right), \quad (5.33)$$

and for holes as

$$p = N_v \exp\left(\frac{F_v - E_p}{kT_h}\right). \quad (5.34)$$

The effective density of states also varies as a function of temperature for the conduction band

$$N_c(T_e) = 2 \left( \frac{2 \pi m_n^* k T_e}{h^2} \right)^{3/2}, \quad (5.35)$$

and the valance band

$$N_p(T_h) = 2 \left( \frac{2 \pi m_p^* k T_v}{h^2} \right)^{3/2}. \quad (5.36)$$

#### 5.8.4 Confined carriers

The QW covers a few nanometres vertically. However, due to the very fast carrier scattering rates (femtoseconds), it is unrealistic to have a carrier temperature profile across the QW. The lattice temperature is averaged over the width of the QW to give a single QW temperature,

$$T_{qw} = \frac{1}{L_{qw}} \int_0^{L_{qw}} T_l(y) dy. \quad (5.37)$$

In this model, the carrier temperature is assumed to be the same as lattice temperature ( $T_L$ ). (This assumption is investigated in detail later.) When parabolic bands are used to simulate the carrier densities, the following analytical formulae are used

$$n_w(F_{nw}, T_{qw}) = k T_{qw} \sum_{ic=1}^{N_{cb}} \rho_{ic}^{2D} \ln(1 + \exp[(F_{nw} - E_{(c,ic)})/k T_{qw}]) \quad , \quad (5.38)$$

$$p_w(F_{pw}, T_{qw}) = k T_{qw} \sum_{iv=1}^{N_{vb}} \rho_{iv}^{2D} \ln(1 + \exp[(F_{pw} - E_{(c,iv)})/k T_{qw}]) \quad . \quad (5.39)$$

These vary as a function of the average QW temperature. When the assumption of parabolic bands is not used, the electron and densities in the QW are calculated using,

$$n(F_e, T_{qw}) = \sum_{i=1}^{N_c} \int_{E_{min}}^{E_{top}} \rho_i^e(E) f_e(F_e, T_{qw}) dE \quad (5.40)$$

and

$$p(F_h, T_{qw}) = \sum_{i=1}^{N_v} \int_{E_{min}}^{E_{top}} \rho_i^h(E) [1 - f_h(F_h, T_{qw})] dE \quad , \quad (5.41)$$

where  $N_{cb/vb}$  is the number of conduction/valance bands in the QW and  $\rho_{e/h}(E)$  is the 2D energy density of states in the conduction/valance band. It is a numerical function calculated from the band structure and includes a factor of two to account for spin degeneracy. The total 2D carrier density is divided by  $L_w$  to give the QW carrier density in  $m^{-3}$ .

The process of calculating the carrier density from the band structure, quasi-Fermi level and temperature requires one integration per subband (equations 5.40-5.41). Numerical integration is a slow process. For this reason, the carrier density is only calculated once and stored to disk in a look up table as a function of carrier temperature and quasi-Fermi level. At simulation time, the simulator reads the table and performs a 2D interpolation to calculate the correct carrier density. Two look up

tables are calculated, one for the electron population and one for the hole population.

### 5.8.5 The intrinsic carrier density

The intrinsic carrier density must be known to calculate the SRH and Auger recombination rates. The intrinsic carrier density is a function of temperature. Using Maxwell-Boltzmann statistics  $n_i(T)$  is given as

$$n_i(T) = \sqrt{N_c N_v} e^{-E_g/2kT} . \quad (5.42)$$

### 5.8.6 Low-field mobility model

The low-field mobility model proposed in [3] is also used in this work. The low-field mobility is given by

$$\mu_{LF}(N, T) = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + (N/N_{ref})^\alpha} , \quad (5.43)$$

where  $\mu_{min}$  is the minimum mobility,  $\mu_{max}$  is the maximum mobility,  $N$  is the total doping density (i.e.  $N_a + N_d$ ),  $N_{ref}$  is a reference doping level and  $\alpha$  is a constant. All of these parameters can vary as a function of temperature, so the following model is used to model their temperature dependence

$$\mu(T) = \mu_0 \left( \frac{T}{300} \right)^{\delta_0} . \quad (5.44)$$

When the average drift velocity is far below the Brownian velocity, the influence of the field on the mobility may be neglected [4], as in this case. A table of parameters may be found in [4].

### 5.8.7 Lattice thermal conductivity

The lattice thermal conductivity in semiconductor materials varies as a strong function of temperature and alloy content. Near room temperature, the variation of the thermal conductivity of a binary compound can be expressed as

$$\kappa_L(T) = \kappa_L(300\text{K}) \left( \frac{T}{T_0} \right)^\alpha, \quad (5.45)$$

where  $T$  is the temperature,  $T_0$  is a normalising constant,  $k_L(x,y)$  is the thermal conductivity and  $\alpha$  describes the temperature dependence of the material. Table 5.1 lists the thermal conductivities for several binary alloys and metals.

Material	$\kappa_L$ (W/Kcm)	$T_0$	$\delta k$	Reference
GaAs	46	300.0	-1.25	[5]
InP	68	300.0	-1.4	[5]
AlAs	80	300.0	-1.37	[5]
InAs	27.3	300	-1.1	[5]
GaP	77	300	-1.4	[5]
AlN	285	300	-1.577	[5]
SiC-6H	490	300	-1.61	[5]
Al <sub>2</sub> O <sub>3</sub>	28	300	-1.0	[5]
SiO <sub>2</sub>	1.38	300	0.33	[5]
Cu	400.1	300.0	-	[6]
Au	316.9	300.0	-	[6]
Polyethylene	0.0017	300.0	-	[7]
BCB	0.2	300	-	[5]

Table 5.1: Lattice thermal conductivities for binary alloys and other common materials used in the simulator.

The thermal conductivity for ternary alloys can be calculated using the binary values and a bowing parameter (see table 5.2)

$$\kappa_{300}^{AB}(x) = \left( \frac{1-x}{\kappa_{300}^A} + \frac{x}{\kappa_{300}^B} + \frac{(1-x) \cdot x}{C_{ABC}} \right)^{-1} . \quad (5.46)$$

The temperature dependence of the thermal conductivity is interpolated between the binary alloy elements using the expression

$$\alpha^{AB} = (1-x) \cdot \alpha^A + x \cdot \alpha^B . \quad (5.47)$$

Material	$C_k$	Reference
AlGaAs	3.3	[5]
InGaAs	1.4	[5]
InAlAs	3.3	[5]
InAsP	3.3	[5]
GaAsP	1.4	[5]
InGaP	1.4	[5]

*Table 5.2: Bowing parameters used in calculating thermal conductivities of common materials used in optoelectronics.*

A similar, but slightly less complete model has been proposed in [4].

### 5.8.8 Specific heat capacity

Specific heat capacity is taken as a polynomial without any temperature dependence, values for common materials used in optoelectronic devices are given in table 5.3.

Material	$C_{300}(J/K\ kg)$	Reference
GaAs	322	[5]
AlAs	441	[5]
InAs	394	[5]
InP	410	[5]

Table 5.3: Specific heat capacities for common materials used in optoelectronic devices.

The following formula was used to interpolate between the binary materials

$$c_L^{AB} = (1-x) \cdot c_L^A + x \cdot c_L^B \quad . \quad (5.48)$$

### 5.8.9 Material density

In transient simulations the material densities are required to calculate the heat stored per unit volume. This is used in the left hand side of equation 5.6. The values for density were taken from [8]

Material	$\rho_{300}(g/cm^3)$	Reference
GaAs	5.32	[8]
AlAs	3.76	[8]
InAs	5.667	[8]
InP	4.81	[8]

Table 5.4: Material densities for common materials used in optoelectronics.

Vegard's law is assumed to hold. Thus, the following formula was used to interpolate between the binary materials

$$\rho_L^{AB} = (1-x) \cdot \rho_L^A + x \cdot \rho_L^B \quad . \quad (5.49)$$

### 5.8.10 Elastic constants ( $C_{11}, C_{12}, C_{44}$ )

The elastic constants ( $C_{11}, C_{12}, C_{44}$ ) are used in the more advanced thermal models (described later) and are also used in calculating material strain. The values of  $C_{11}, C_{12}$  and  $C_{44}$  are taken from [9].

Material	$C_{11} \times 10^{10}$ (N/m <sup>2</sup> )	$C_{12} \times 10^{10}$ (N/m <sup>2</sup> )	$C_{44} \times 10^{10}$ (N/m <sup>2</sup> )	Ref
$Al_xGa_{(1-x)}As$	$11.88 + 0.14x$	$5.38 + 0.32x$	$5.94 - 0.05x$	[10]
$InP$	10.11	5.61	4.56	[11]

Table 5.5: Polynomials describing elastic constants

### 5.8.11 Other parameters

Some other material parameters used within the thermal model (either within this chapter or in the more advanced thermal models in later chapters) were introduced into the database and described using polynomials of the form

$$Z(x, y) = \sum_{i=0}^Y \sum_{j=0}^X C_{ij} x^i y^j, \quad (5.50)$$

where  $C_{ij}$  are constants,  $x_i$  and  $y_j$  material compositions raised to the power of  $i$  and  $j$  respectively. These parameters include:

- Average atomic spacing
- Relative atomic mass
- The number of atoms per lattice base

### 5.8.12 Bulk spontaneous emission

Spontaneous emission is the process by which an electron in the conduction band

recombines with a hole in the valence band, emitting a photon. This process is described in more detail in Chapter 2. In bulk semiconductors, the rate of this process is described by

$$R^{Spont} = B(np - n_i^2) \quad , \quad (5.51)$$

where (in steady-state)  $B$  is the capture rate [12]. Although  $B$  is not taken to be a temperature dependent parameter,  $R^{Spont}$  is affected by the temperature dependence of  $n$ ,  $p$  and  $n_i$ . For bulk materials,  $R^{Spont}$  is very low, since the  $np$  term in equation 5.51 is small. The value used in bulk AlGaAs is given in table 5.6.

Capture Rate	Value ( $\text{m}^3\text{s}^{-1}$ )	Reference
$B$	$1.7 \times 10^{-16}$	[13]

Table 5.6: Bulk Spontaneous Emission.

### 5.8.13 Bulk Shockley-Read-Hall (SRH)

SRH recombination describes the recombination of electrons and holes at deep level defects. The energy is released to the lattice as phonons. The equation describing this process is given by

$$R^{SRH} = \frac{(np - n_i^2)}{\tau_{p0}(n + n_i) + \tau_{n0}(p + n_i)} \quad , \quad (5.52)$$

The capture rates  $\tau_{p0}$  and  $\tau_{n0}$  used are given in table 5.7.

Although the capture rates are not temperature dependant, the overall rate of SRH

Capture Rate	Value (ns)	Reference
$\tau_{p0}$	10	[14]
$\tau_{n0}$	10	[14]

Table 5.7: SRH Capture rates used in the device simulator.

recombination is temperature dependant. In high quality materials such as those required for lasers, bulk SRH recombination rates are low.

#### 5.8.14 Bulk Auger recombination

The Auger recombination rate is given by [4]

$$R^{Auger} = (C_n n + C_p p)(np - n_i^2) \quad . \quad (5.53)$$

The Auger recombination coefficients are parametrised as temperature dependant parameters using [15,16]

$$C_{n,p} = C_0 \exp \left[ \frac{E_0}{k_B} \left( \frac{1}{T} - \frac{1}{T_0} \right) \right] \quad . \quad (5.54)$$

The coefficients  $C_0$ ,  $E_0$  and  $T_0$  are given in table 5.8. These were obtained from numerical fits to experimental and theoretical calculations.

Material system	$C_{n,p0}$ ( $m^6 s^{-1}$ )	$E_0$	$T_0$
GaAlAs	$4.22 \times 10^{-42}$ [15]	0.0404	300.0

Table 5.8: The Auger recombination coefficients used within the device simulator.

#### 5.8.15 Auger and SRH recombination in the quantum well

It is usual to fit the simulated L-I curves to experimental data by varying the QW Auger and SRH recombination constants. Auger recombination in the QW is described by

$$R_{qw}^{Auger} = (C_n^{qw} n + C_p^{qw} p)(np - n_i^2) \quad (5.55)$$

where

$$C_n = 1 \times 10^{-41} \text{ m}^6 \text{ s} \quad (5.56)$$

and

$$C_p^{qw}(T) = C_{p0}^{qw} [1 + \beta(T - T_0)^n] \quad (5.57)$$

with  $C_0 = 1 \times 10^{-28} \text{ m}^6 \text{ s}$ ,  $\beta = 5 \times 10^{-4}$  and  $n = 2.05$ ,  $T_0 = 350 \text{ K}$ . The SRH recombination rate for the QW is given by

$$R^{SRH} = \frac{(np - n_i^2)}{\tau_{p0}(n + n_i) + \tau_{n0}(p + n_i)} \quad (5.58)$$

The capture times used are given in table 5.9.

Capture Rate	Value (ns)
$\tau_{p0}$	6
$\tau_{np}$	6

*Table 5.9: SRH Capture rates used in the device simulator.*

The fitting to experiment was performed by J. J. Lim as part of the EU-Project FAST-ACCESS (IST-004772).

### 5.8.16 Maximum temperature

If a device is simulated far into roll-over, very high device temperatures can be

calculated ( $T_L > 1000\text{K}$ ). At such high temperatures, the device would quickly degrade and die through defect formation. This regime is of interest when investigating effects such as catastrophic optical damage (COD). It is, however, not of immediate interest in every day device design and optimisation. Furthermore, although the laser solver is stable enough to solve the problem during roll-over, it requires many iterations and the material models may not hold – thus bringing into question the validity of the results. Therefore, the melting point of gold (1337K) [17], which is commonly used to form the top contact and wire bond was set as the maximum simulation temperature. If such a temperature is calculated at any point within the mesh, the simulation is stopped. Thus, no computational time is wasted on solving for temperatures far outside the devices operating regime.

## **5.9 The carrier transport model**

### **5.9.1 Thermodynamic treatments and hydrodynamic models**

The two key approaches used to derive the heating terms related to current flux within devices are the thermodynamic and hydrodynamic approaches. The thermodynamic approach was originally proposed by Wachutka [18-20] and is based on irreversible thermodynamics. The model takes as its starting point the phenomenological current relations and makes use of Onsager's relations to describe the heat flux densities along with the principle of local conservation of energy. The model describes a system close to equilibrium, where Onsager's relations and Maxwell-Boltzmann statistics are valid. The model also depends on the correct evaluation of the thermoelectric powers for electrons and holes.

The hydrodynamic model [21] takes as its starting point the Boltzmann Transport equation (BTE). Many methods have been proposed for solving the BTE. The Monte-Carlo method provides the most detailed solution method, but is computationally too slow for day-to-day device design and optimisation [22]. In the hydrodynamic model, a moment expansion of the BTE is taken, giving expressions for carrier continuity and conservation of momentum and energy, which are derived from first principles. This approach has been widely used for decades in semiconductor modelling (e.g. MINIMOS [23]). The model is more readily extendable to FD statistics and for the inclusion of vector temperatures [24], tensor masses [24] and multiple sub bands. Thus, the moment expansion of the BTE is more suitable for systems far away from thermal equilibrium, as is often the case for laser diodes. Under certain conditions (e.g. close to equilibrium), the equivalence of the thermodynamic and hydrodynamic approaches has been demonstrated [25].

### **5.9.2 The Drift Diffusion (DD) and Energy Balance (EB) models**

R. Stratton [21] first performed a moment expansion on the BTE, deriving terms to describe the carrier and energy transport in a semiconductor. Later, K. Bløtekjær [26] extended the theory to two valley semiconductors. The work of Azoff [22] extended the hydrodynamic model to include inhomogeneous semiconductors. More recent models have introduced non-parabolicity of the band structure, and higher moments up to and including the 6<sup>th</sup> moment [27]. Anisotropy in the temperature distribution has been introduced into some device simulators and it has been shown that the

temperature distribution is elongated along the direction of the field [24].

Each moment of the expansion of the BTE requires the result of the previous expansion for solution, resulting in an infinite series of moments. The series must be brought to a close with a so called closure relation. In models which use the first four moments (0-3<sup>rd</sup> moments), it is common to close the series with the heat flux generated by the gradient of the carrier temperature profile,

$$\mathbf{Q} = -\kappa \nabla T_{elh} . \quad (5.59)$$

The validity of this relation has been questioned in the literature [25-27]. A review of closure relations and comparison to full Monte-Carlo simulation can be found in [24].

The BTE is given as

$$\frac{\partial f}{\partial t} + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla f + \frac{\partial \mathbf{p}}{\partial t} \cdot \nabla_{\mathbf{p}} f = C[f] , \quad (5.60)$$

where  $f$  is the carrier distribution function,  $\mathbf{r}$  is the position vector,  $\mathbf{p}$  is the momentum vector,  $\nabla$  is the gradient operator with respect to position,  $\nabla_{\mathbf{k}}$  is the gradient operator with respect to momentum and  $C[f]$  accounts for carrier scattering events (usually approximated by the relaxation time approximation). This is a 7 dimensional integro-differential equation. In this work, we follow the work of Azzoff [22]. By multiplying the BTE by 1,  $mv_i$  and  $m|v|^2/2$  and integrating over  $k$ -space the carrier continuity, momentum conservation and energy conservation equations can be

obtained respectively. This method is fully described in appendix A, but the outline of the method and key results are given below.

The 0<sup>th</sup> moment of the BTE is calculated by multiplying equation 5.60 by 1 and integrating over  $k$ -space. In this way, the the current continuity equations for electrons and holes can be obtained

$$\nabla \cdot \mathbf{J}_n = q \cdot \left( R + \frac{\partial n}{\partial t} \right) , \quad (5.61)$$

$$\nabla \cdot \mathbf{J}_p = q \cdot \left( R + \frac{\partial p}{\partial t} \right) , \quad (5.62)$$

where  $R$  is the carrier recombination rate and  $J_n, J_p$  are the electron and hole fluxes, respectively.

The 1<sup>st</sup> moment of the BTE can be derived by multiplying equation 5.60 by momentum ( $m\mathbf{v}$ ) and integrating over  $k$ -space. This gives the drift diffusion equations for electrons and holes

$$\mathbf{J}_n = \mu_e n \left[ \nabla E_c - \frac{3}{2} k T_e \nabla \ln m_e^* + k \nabla T_e \right] + k T_e \mu_e \nabla n , \quad (5.63)$$

$$\mathbf{J}_p = \mu_h p \left[ \nabla E_v - \frac{3}{2} k T_h \nabla \ln m_h^* - k \nabla T_h \right] - k T_h \mu_h \nabla p . \quad (5.64)$$

The 2<sup>nd</sup> moment of the BTE can be derived by multiplying equation 5.60 by energy

( $mlv|v|^2$ ) and integrating over  $k$ -space. This gives the energy conservation equations [28] for electrons and holes

$$\nabla \cdot \mathbf{S}_n = \nabla \left( \frac{E_c}{q} - \Phi \right) \cdot \mathbf{J}_n - \frac{3k_B}{2} \cdot \left( \frac{\partial(nT_n)}{\partial t} + RT_n + n \frac{T_n - T_L}{\tau_n} \right), \quad (5.65)$$

$$\nabla \cdot \mathbf{S}_p = \nabla \left( \frac{E_v}{q} - \Phi \right) \cdot \mathbf{J}_p - \frac{3k_B}{2} \cdot \left( \frac{\partial(pT_p)}{\partial t} + RT_p + p \frac{T_p - T_L}{\tau_p} \right), \quad (5.66)$$

where,

$$\mathbf{S}_p = -\kappa_p \nabla T_p + \frac{5}{2} \frac{k_B T_p}{q} \mathbf{J}_p, \quad (5.67)$$

$$\text{and } \mathbf{S}_n = -\kappa_n \nabla T_n + \frac{5}{2} \frac{k_B T_n}{q} \mathbf{J}_n \quad (5.68)$$

are the carrier energy fluxes and

$$\kappa_n = \left( \frac{5}{2} + c_p \right) \frac{k_B^2}{q} T_e \mu_e n, \quad (5.69)$$

$$\text{and } \kappa_p = \left( \frac{5}{2} + c_p \right) \frac{k_B^2}{q} T_p \mu_p p \quad (5.70)$$

are the carrier thermal conductivities.  $T_n$  is defined as the electron carrier temperature,  $T_L$  as the lattice temperature,  $T_p$  as the hole carrier temperature,  $\tau_n$  as the electron relaxation time,  $\tau_p$  as the hole relaxation time and  $\Phi$  as the electrostatic potential.

In the hydrodynamic model, the energy lost by the carriers is given to the lattice.

Thus, the bulk lattice heating equation is

$$H = \frac{3k_B}{2} \cdot \left( n \frac{T_n - T_L}{\tau_n} + p \frac{T_p - T_L}{\tau_p} \right) + R(E_c - E_v), \quad (5.71)$$

where  $R(E_c-E_v)$  is the recombination multiplied by the band gap (i.e. the dark carrier recombination term) if the electron, hole and lattice temperatures  $T_e=T_h=T_l$  are assumed to be the same, and the lattice scattering is assumed to be instantaneous, then the lattice heat equation may be written as

$$H = \nabla \left( \frac{E_c}{q} - \phi \right) \cdot \mathbf{J}_n + \nabla \left( \frac{E_v}{q} - \phi \right) \cdot \mathbf{J}_p + R(E_c - E_v) \quad . \quad (5.72)$$

For computational speed and stability, the drift diffusion (DD) model will be used. Therefore, 5.61-5.64 will be used along with 5.72 to give the DD model with the lattice heating model.

### 5.9.3 Scharfetter-Gummel discretisation of the drift diffusion equations

The carrier continuity equations (both in the bulk (5.73 and 5.74) and in the QW (5.75 and 5.76)) require the derivative of the carrier fluxes to be evaluated at each mesh point,

$$\nabla \cdot \mathbf{J}_n - qR = 0 \quad (5.73)$$

$$\nabla \cdot \mathbf{J}_p - qR = 0 \quad (5.74)$$

$$\nabla \cdot \mathbf{J}_{nqw} - qR = 0 \quad (5.75)$$

$$\nabla \cdot \mathbf{J}_{pqw} - qR = 0 \quad (5.76)$$

To do this, the carrier fluxes must be known mid way between the mesh points

$$\frac{\partial J_i}{\partial x} = \frac{J_{i+1/2} - J_{i-1/2}}{\Delta x} \quad . \quad (5.77)$$

Unfortunately, the carrier density varies exponentially as a function of potential. Thus, when calculating the current flux at the mid point  $i \pm 1/2$  it is not possible to simply average the current fluxes at  $J_i$  and  $J_{i \pm 1}$  to obtain an accurate value for  $J_{i \pm 1/2}$ . It has been shown necessary to use the Scharfetter-Gummel [29] scheme to produce stable, well convergent solutions. In the following section, the temperature dependent drift diffusion equations are discretised using the Scharfetter-Gummel approach. The temperature dependent drift diffusion equation for electrons

$$J_n = \mu_e n \left[ \nabla E_c - \frac{3}{2} k T_e \nabla \ln m_e^* + k \nabla T_e \right] + k T_e \mu_e \nabla n \quad (5.78)$$

can be rewritten as

$$\frac{J_n}{k T_e \mu_e} = n \left[ \frac{\nabla E_c}{k T_e} - \frac{3}{2} \nabla \ln m_e^* + \frac{\nabla T_e}{T_e} \right] + \nabla n \quad , \quad (5.79)$$

which is a first order differential equation of the form

$$C = A n + \frac{\partial n}{\partial x} \quad . \quad (5.80)$$

In 1D, the constants can be defined as

$$C = \frac{J_n}{k T_e \mu_e} \quad \text{and} \quad (5.81)$$

$$A = \frac{1}{k T_e} \left[ \frac{\partial E_c}{\partial x} - \frac{3}{2} \frac{k T_e}{m_e^*} \frac{\partial m_e^*}{\partial x} + k \frac{\partial T_e}{\partial x} \right] \quad . \quad (5.82)$$

Equation 5.80 may be rewritten as

$$\frac{\partial}{\partial x} (n e^{Ax}) = C e^{Ax} \quad . \quad (5.83)$$

This may be solved by integration between the mesh points  $i$  and  $i+1$

$$e^{Ax_{i+1}} n_{i+1} - e^{Ax_i} n_i = \frac{C}{A} (e^{Ax_{i+1}} - e^{Ax_i}) \quad . \quad (5.84)$$

Rearranging gives,

$$C = A \frac{e^{Ax_{i+1}} n_{i+1} - e^{Ax_i} n_i}{e^{Ax_{i+1}} - e^{Ax_i}} \quad (5.85)$$

$$C = \frac{Ae^{Ax_{i+1}} n_{i+1}}{e^{Ax_{i+1}} - e^{Ax_i}} - \frac{Ae^{Ax_i} n_i}{e^{Ax_{i+1}} - e^{Ax_i}} \quad (5.86)$$

and using

$$h_{i+1/2} = x_{i+1} - x_i \quad (5.87)$$

The constant  $C$  may be written as

$$C = \frac{An_{i+1}}{1 - e^{Ax_i - Ax_{i+1}}} - \frac{An_i}{e^{Ax_{i+1} - Ax_i} - 1} \quad . \quad (5.88)$$

Using the Bernoulli function [1]

$$B(x) = \frac{x}{e^x - 1} \quad (5.89)$$

and rewriting

$$C_{i+1/2} = \frac{1}{h_{i+1/2}} \left[ B(-Ah_{i+1/2}) n_{i+1} - B(Ah_{i+1/2}) n_i \right]_{i+1/2} \quad (5.90)$$

Finally, defining  $\alpha$  as

$$\alpha = Ah_{i+1/2} = \frac{h_{i+1/2}}{kT} \left[ \frac{\partial E_c}{\partial x} - \frac{3}{2} \frac{kT}{m_e^*} \frac{\partial m_e^*}{\partial x} + k \frac{\partial T}{\partial x} \right]_{i+1/2} \quad , \quad (5.91)$$

and substituting 5.81 results in the current flux at the mid point ( $i+1/2$ ) of the mesh expressed in terms of the carrier densities on the mesh points

$$J_{i+1/2} = \frac{kT\mu_e}{h_{i+1/2}} \left[ B(-\alpha)n_{i+1} - B(\alpha)n_i \right] \Big|_{i+1/2} . \quad (5.92)$$

Discretising  $\alpha$  gives

$$\alpha = 2 \left[ \frac{1}{k} \frac{E_{c_{i+1}} - E_{c_i}}{T_{i+1} + T_i} - \frac{3}{2} \frac{m_{e_{i+1}}^* - m_{e_i}^*}{m_{e_{i+1}}^* + m_{e_i}^*} + \frac{T_{i+1} - T_i}{T_{i+1} + T_i} \right] . \quad (5.93)$$

It is common in the literature to replace the  $m_e^*$  terms with the 3D effective density of states [31] using the following procedure

$$\frac{3}{2} \frac{1}{m^*} \frac{\partial m^*}{\partial x} = \frac{3}{2} \frac{\partial \ln m^*}{\partial x} = \frac{\partial \ln m^{*3/2}}{\partial x} \quad (5.94)$$

then substituting the effective mass term with the effective density of states,

$$N_c = 2 \left( \frac{2\pi kT_e}{h^2} \right)^{3/2} m_n^{*3/2} , \quad (5.95)$$

which gives

$$\frac{1}{m^{*3/2}} \frac{\partial m^{*3/2}}{\partial x} = \frac{1}{N_c} \frac{\partial N_c}{\partial x} . \quad (5.96)$$

However, this procedure assumes that the temperature term in 5.95 is independent of position. Thus, 5.96 does not hold for the thermal case - this is often over looked.

The same procedure can be repeated for the drift diffusion equation for holes.

$$J_p = \mu_h p \left[ \nabla E_v - \frac{3}{2} kT_h \nabla \ln m_h^* - k \nabla T_h \right] - kT_h \mu_h \nabla p , \quad (5.97)$$

When discretised using the above method, this gives the current at  $i+1/2$  as

$$J_{h_{i+1/2}} = \frac{k T_h \mu_h}{h_{i+1/2}} \left[ B(\alpha_{i+1/2}) n_{i+1} - B(-\alpha_{i+1/2}) n_i \right]_{i+1/2}, \quad (5.98)$$

with

$$\alpha = 2 \left[ \frac{1}{k} \frac{E_{v_{i+1}} - E_{v_i}}{T_{i+1} - T_i} - \frac{3}{2} \frac{m_{v_{i+1}} - m_{v_i}}{m_{v_{i+1}} + m_{v_i}} - \frac{T_{v_{i+1}} - T_{v_i}}{T_{v_{i+1}} + T_{v_i}} \right]. \quad (5.99)$$

## 5.10 Gain

### 5.10.1 Calculation of transition rate

As the carrier temperature increases, the FD distribution function spreads, resulting in fewer carriers in the lasing states and thus reduced gain. This is one of the main causes of thermal roll-over in laser diodes. It is therefore essential to have the correct dependence of gain on temperature for accurate device modelling. A program for the calculation of gain and band structure was already available, but had to be modified to generate temperature dependant gain tables. The laser simulator also had to be modified to accept the tables. The key physical equation used to describe the emission of photons from the semiconductor is Fermi's Golden rule. This gives the transition rate ( $s^{-1}$ ) [30] as

$$W_{e \rightarrow h} = \frac{2\pi}{\hbar} |H'_{eh}|^2 \delta(E_e - E_h - \hbar\omega), \quad (5.100)$$

where  $H'_{eh}$  is the matrix element [30],  $E_e$  is the electron energy,  $E_h$  is the hole energy and  $\hbar\omega$  is the energy of the interacting photon. The matrix element in bra-ket (or Dirac) notation is given by

$$H'_{eh} = \langle \Psi_h | H'(r) | \Psi_e \rangle = \int_V \Psi_h^* H'(r) \Psi_e d^3r, \quad (5.101)$$

where the electron and hole wave functions are given by  $\Psi_e$  and  $\Psi_h$  respectively. The

perturbation Hamiltonian is defined as [30]

$$H'(\mathbf{r}) \equiv \frac{e}{2m_0} A(\mathbf{r}) \hat{\mathbf{e}} \cdot \mathbf{p} \quad . \quad (5.102)$$

After some manipulation [30],  $H'_{eh}$  may be written as

$$|H'_{eh}|^2 = \left( \frac{eA_0}{2m_0} \right)^2 |M_T|^2 \quad , \quad (5.103)$$

where for transverse magnetic (TM) polarisations

$$|M_T|^2 = |M|^2 [4/3 |\langle F_l | F_e \rangle|^2] \quad , \quad (5.104)$$

where  $F_l$  and  $F_e$  are the light hole and conduction band envelope functions respectively. For the (TE) polarisations,  $M_T$  is given as

$$|M_T|^2 = |M|^2 [|\langle F_h | F_e \rangle|^2 + 1/3 |\langle F_l | F_e \rangle|^2] \quad , \quad (5.105)$$

where  $F_l$  is the heavy hole envelope function. Values of  $M$  can be determined experimentally. Where  $F_l=0$  for heavy hole states and  $F_h=0$  for light hole states. The transition rate per unit volume is calculated by integrating over the number of allowed electron-hole transitions, then dividing by the unit volume in real space

$$W_{c \rightarrow v} = \frac{1}{V} \int W_{e \rightarrow h} dN_s = \int W_{w \rightarrow h} \frac{1}{V} \frac{dN_s}{dk} dk \quad (5.106)$$

giving the number of transitions per unit volume per second ( $s^{-1}m^{-3}$ ). Rewriting using Fermi's golden rule and the density of states as defined in Chapter 2 gives

$$W_{c \rightarrow v} = \frac{2\pi}{\hbar} \int |H'_{eh}| \delta(E_{eh} - \hbar\omega) \rho(k) dk \quad , \quad (5.107)$$

where  $E_{eh}=E_c-E_h$  has been defined as the energy separation of the interacting carrier pair. The integration is performed by a change in variables  $x=E_{eh}-\hbar\omega$

$$W_{c \rightarrow v} = \frac{2\pi}{\hbar} \int |H'_{eh}|^2 \delta(x) \rho(k) \frac{dx}{(dE_{eh}/dk)} \quad (5.108)$$

and noting the delta function results in

$$W_{c \rightarrow v} = \frac{2\pi}{\hbar} |H'_{eh}|^2 \left[ \frac{\rho(k)}{(dE_{eh}/dk)} \right]_{E_{eh}=\hbar\omega} \quad (5.109)$$

The reduced density of states is then defined as

$$\rho_{red} = \left[ \frac{\rho(k)}{(dE_{eh}/dk)} \right] \quad (5.110)$$

By rewriting equation 5.109 using equation 5.110 and then multiplying by the probability of the initial state containing an electron ( $f_c$ ), then by the probability of the final state being empty ( $1-f_v$ ) results in

$$W_{c \rightarrow v} = \frac{2\pi}{\hbar} |H'_{eh}|^2 \rho_{red} f_c (1-f_v) \quad (5.111)$$

### 5.10.2 Calculation of material gain

Gain is defined as the fractional increase in photon flux per unit length [30] and can be written by

$$g \equiv \frac{1}{\Phi} \frac{d\Phi}{dz} = \frac{W_{c \rightarrow v} - W_{v \rightarrow c}}{\Phi} \quad (5.112)$$

where  $\Phi$  is the photon flux.  $\Phi$  is given by

$$\Phi(\omega) = \frac{1}{\hbar\omega} \left( \frac{c}{n_g} \right) \left( \frac{1}{2} n n_g \epsilon \omega^2 A_0^2 \right) \quad [30], \quad (5.113)$$

where  $c$  is the speed of light,  $n$  is the refractive index,  $n_g$  is the group index and  $\epsilon$  is the permittivity of free space. Substituting equations 5.111, 5.103 and 5.113 into equation 5.112 gives equation 5.114, an expression for the material gain

$$g_{\text{sub}}(\hbar\omega) = \left( \frac{1}{\hbar\omega} \right) \frac{\pi e^2 \hbar}{\epsilon_0 c m_0^2 \bar{n}} |M_T|^2 \rho_{\text{red}}(E_{eh} - E'_g)(f_c - f_v) . \quad (5.114)$$

The total material gain is given by a summation over all subbands

$$g(\hbar\omega) = \sum_{n_c} \sum_{n_v} g_{\text{sub}}(\hbar\omega, n_c, n_v) . \quad (5.115)$$

Due to the finite lifetime of the carriers in these states, the gain spectrum must be broadened with an appropriate broadening function. The Lorentzian function

$$L(E_{eh}) = \frac{1}{\pi} \frac{\hbar/\tau}{(E_{eh} - \hbar\omega)^2 + (\hbar/\tau)^2} \quad (5.116)$$

is often used for this purpose, where  $\tau$  is the broadening lifetime. Equation 5.116 is convolved with the gain spectra to give the broadened gain spectra using

$$G(\hbar\omega) = \int g(\hbar\omega) L(E_{eh}) dE_{eh} . \quad (5.117)$$

As was done for the carrier densities, the material gain was tabulated as a function of wavelength, electron density, hole density and temperature.

### 5.10.3 Spontaneous emission

The downward transition probability is proportional to  $(\langle n_{\text{ph}} \rangle + 1)\hbar\omega$  where  $\langle n_{\text{ph}} \rangle$  is the number of photons in a mode. Even if no photons [30] are present in the mode, there is still a downward transition probability. This is the reason for spontaneous emission. It can be shown that the spontaneous emission rate is given by [30]

$$R_{sp}(\hbar\omega) = \left( \frac{1}{\hbar\omega} \right) \frac{\pi e^2 \hbar}{\bar{n} \bar{n}_g \epsilon_0 m_0^2} |M_{ave}|^2 \rho_{\text{red}}(E_{eh} - E'_g) \rho_{\text{opt}} f_c (1 - f_v) , \quad (5.118)$$

where

$$|M_{ave}|^2 \equiv \frac{1}{3} \sum_{i=0}^3 |M_T|^2 . \quad (5.119)$$

This function is dependent upon temperature and quasi-Fermi levels. It was tabulated in a three dimensional look up table for use within the simulations.

## 5.11 Bulk heat generation

### 5.11.1 Average separation of carriers

In device models used to simulate silicon devices, it is common to assume that recombination occurs between the conduction and valance band edges,  $E_c$  and  $E_v$ , respectively [31]. This approach assumes that all of the electrons in the conduction band are at energy  $E_c$  and that all of the holes in the valence band are at energy  $E_v$ . Consequently, this neglects the thermal energy of the carriers. Admittedly, this is a small energy when one compares the band gap energy and the thermal energy. In high quality silicon devices where recombination rates are small, the consequences of neglecting the thermal energy are relatively small. However, in lasers where recombination is the dominant process, the thermal energy of a bulk carrier pair at 300K is 0.078eV and the band gap is around 0.91eV. Thus, the thermal energy of the carriers accounts for around 8% of the recombination energy and must be taken into account.

The average energy of a carrier is given by

$$\bar{E} = \frac{\int_{E_{min}}^{\infty} E \rho(E)_{3D} f(E) dE}{\int_{E_{min}}^{\infty} \rho(E)_{3D} f(E) dE} \quad (5.120)$$

The density of 3D states assuming parabolic bands is given as

$$\rho(E)_{3D} = \frac{\sqrt{E}}{4\pi^2} \left( \frac{2m^*}{\hbar^2} \right)^{3/2} \quad (5.121)$$

$f(E)$  is the occupational probability given by Fermi-Dirac statistics as

$$f(E) = \frac{1}{1 + e^{(E-E_f)/kT}} \quad (5.122)$$

If full FD stastics are used, equation 5.120 must be evaluated numerically. However, when Maxwell-Boltzmann statistics are assumed for the bulk carrier distributions,  $f(E)$  can be approximated as

$$f(E) = e^{-(E-E_f)/kT} \quad (5.123)$$

The denominator of equation 5.120 simply gives the carrier density and can be written as  $n$ . The numerator can be evaluated by integrating by parts.

$$\begin{aligned} & \frac{1}{4\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \int_{E_{min}}^{\infty} E^{3/2} e^{(E_f-E)/kT} dE \quad (5.124) \\ &= \frac{1}{4\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \left[ -kT \cdot E^{3/2} e^{(E_f-E)/kT} \Big|_0^{\infty} + \frac{3}{2} kT \int_0^{\infty} e^{(E_f-E)/kT} E^{1/2} dE \right] \end{aligned}$$

(5.125)

$$= \frac{1}{4\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \left[ 0 + \frac{3}{2} kT \int_0^\infty e^{(E_f - E)/kT} E^{1/2} dE \right] \quad (5.126)$$

The integral on the right hand side of equation 5.126 is the carrier density multiplied by  $3/2 kT$ . Thus, the result of the integral is

$$\frac{1}{4\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} \int_{E_{min}}^\infty E^{3/2} e^{(E_f - E)/kT} dE = \frac{3}{2} kT n \quad (5.127)$$

and dividing by the denominator in equation 5.120 gives

$$\bar{E} = \frac{3}{2} kT \quad (5.128)$$

Note that this is the classical result one would expect from thermodynamics using the principle of partition of energy. Therefore, the average energy of a carrier in the conduction band is

$$\bar{E}_c = E_c + \frac{3}{2} kT_e \quad (5.129)$$

and the average energy of a carrier in the valance band is

$$\bar{E}_v = E_v - \frac{3}{2} kT_h \quad (5.130)$$

The average separation of a carrier in terms of energy can be written as

$$\bar{E}_{eh} = \bar{E}_v - \bar{E}_c \quad (5.131)$$

Therefore, the recombination heat source is given as

$$H_{\text{recom bulk}} = R_{\text{bulk}} \cdot \bar{E}_{eh} \quad (5.132)$$

where R is the net rate of recombination in the bulk given by

$$R_{\text{bulk}} = (R_{SRH} + R_{Auger} + R_{spon}) \cdot \bar{E}_{eh} \quad (5.133)$$

### 5.11.2 Free carrier absorption

Free carrier absorption (FCA) is a process whereby a photon is absorbed by an electron or a hole, exciting the carrier from a low energy state to a higher energy state in the same band (conduction or valance). The carrier then relaxes by emitting heat (phonons) to the lattice [32]. The most significant contribution to free carrier absorption usually occurs in the valance band due to the presence of the 3 valance bands (lh, hh and so)

The free carrier absorption loss is given by,

$$\alpha_{fc} = \alpha_n n + \alpha_p p \quad (5.134)$$

where  $\alpha_n$  and  $\alpha_p$  are the free-carrier absorption constants [33]. The constants for the bulk GaAs are given in table 5.10.

Constant	Value (m <sup>2</sup> )
$\alpha_n$	$3 \times 10^{-24}$
$\alpha_p$	$7 \times 10^{-24}$

*Table 5.10: Bulk FCA constants for GaAs [34].*

If the photon density as a function of position is given by  $\Phi_{PH}(x,y)$ , the energy of the light interacting with the carriers is given by  $\hbar\omega$  and the group velocity of the light within the cavity is given by  $v_g$ . Then, the total heat generation per unit volume in the bulk is given by

$$H_{fca} = (\alpha_n n_T + \alpha_p p_T) \Phi_{PH}(x, y) \hbar \omega v_g \quad (5.135)$$

## 5.12 Heating in the quantum well

The vast majority of carriers entering the device recombine within the quantum well. The heating in this region must be accurately treated for a device model to be predictive. The major heat generation mechanisms are free carrier absorption, SRH recombination, Auger recombination, spontaneous emission, lateral Joule heating and heating due to carrier capture and relaxation. In the following section, the heating terms describing carrier heating in the QW will be described.

### 5.12.1 Capture heating

Before a carrier in the QW can undergo stimulated emission, it must first be captured from the 3D bulk states into the 2D confined states. As the carriers relax down the QW, they must lose their kinetic energy and momentum. The polar optical phonon scattering mechanism is the dominant scattering mechanism in GaAs based semiconductors. Thus, the energy and momentum of the carriers are lost via LO-phonon emission. The distance from the average 3D conduction band energy to the bottom of the band edge is given by

$$\Delta E_c = \bar{E}_c - E_c^{qw} \quad , \quad (5.136)$$

where  $E_c^{qw}$  is the bottom of lowest band in the QW. For the valance band, the distance from the 3D-states to the bottom of the lowest band is given by

$$\Delta E_v = \bar{E}_v - E_v^{qw} \quad , \quad (5.137)$$

where  $E_v^{qw}$  is the bottom of lowest band in the valance band QW. Thus, the heat generated due to capture heat is given by

$$H_{cap} = \Delta E_c R_{cap}^e + \Delta E_v R_{cap}^h \quad , \quad (5.138)$$

where  $R_{cap}^e$  and  $R_{cap}^h$  are the net electron and hole QW capture rates. We now define for later use the QW band gap as the distance between the bottom of the lowest conduction subband and top of the highest valence subband

$$E_g^{qw} = E_c^{qw} - E_v^{qw} \quad . \quad (5.139)$$

### 5.12.2 Stimulated emission

No heat is emitted during stimulated emission. However, equation 5.138 overestimates the energy released during carrier relaxation. The carriers do not lase from the bottom of the lowest band in the QW. In fact, they lase from slightly above this energy. Thus, equation 5.138 simulates the carriers involved in stimulated emission relaxing too far. To readjust this energy balance, a small energy equal to the distance between the bottom of the lowest band and the energy at which the carriers recombine through stimulated emission must be subtracted from the total heat generation. This must be done for both the conduction and valence bands in the QW. Equation 5.138 assumes the carriers have fallen  $\Delta E_c$  and  $\Delta E_v$  respectively. Therefore, if stimulated emission were happening at the bottom of the band, the carriers would only have to fall through  $E_g^{qw}$  i.e.  $\bar{E}_g - \Delta E_c - \Delta E_v$  . However, we already know that the energy an electron must fall through to recombine with a hole is  $\hbar\omega_i$  where  $i$  is the mode number. Thus the carriers which recombine via stimulated emission have been calculated as generating  $\hbar\omega_i - E_g^{qw}$  more heat energy than they should have. Therefore, the heat

$$H_{spon} = \sum_{i=0}^N R_{stim}^i (E_g^{qw} - \hbar \omega_i) \quad (5.140)$$

needs to be subtracted from the net heat generation term, where N is the number of modes.

### 5.12.3 Spontaneous emission

Light emitted due to spontaneous emission is radiated in all directions. The spectrum of the light emitted is also very broad compared to the lasing mode, thus very little light escapes the cavity. Most of the light is absorbed and eventually given to the lattice as heat. For this reason, spontaneous emission is modelled as a heat source

$$H_{spon} = E_g^{qw} R^{spon} . \quad (5.141)$$

### 5.12.4 Lateral Joule heating

The QW in a laser diode is usually undoped and thus has a very high mobility. It therefore produces a very small amount of Joule heating. Joule heating due to the lateral movement of carriers within the QW is generally the most insignificant heat source within a device. However, it should be taken into account. The following equation takes accounts for Joule heating due to lateral current flow in the QW

$$H_{Joule} = J_e^{qw} \frac{\partial}{\partial x} E_c + J_h^{qw} \frac{\partial}{\partial x} E_v . \quad (5.142)$$

### 5.12.5 Shockley-Read-Hall (SRH) heating

Shockley-Read-Hall is a dark recombination process, as explained in detail in Chapter

2. The heat generated in the process is given by

$$H_{SRH}^{qw} = R_{SRH} \cdot E_g^{qw} . \quad (5.143)$$

Although equation 5.138 overestimates the heat generated in the capture of the carriers involved in SRH recombination, this does not matter because the heat generated in SRH recombination will be underestimated by the same amount.

### 5.12.6 Free carrier absorption

The heat generated by free carrier absorption is given by

$$H_{fca}^{qw} = (\alpha_n^{QW} n_T + \alpha_p^{QW} p_T) \Phi_{PH}(x, y) \hbar \omega v_g , \quad (5.144)$$

where  $\alpha_n^{QW}$  and  $\alpha_p^{QW}$  are the free carrier absorption constants associated with the QW.

### 5.13 Total lattice heat

The total quantum well lattice heating term is given by

$$H_{qw} = H_{joule}^{qw} + H_{FCA}^{qw} + H_{Aug}^{qw} + H_{SHR}^{qw} . \quad (5.145)$$

Therefore, the total heating term is given as

$$H_{qw} = H_{qw} + H_{recom\ bulk} + H_{fca} + H_{joule}^{bulk} . \quad (5.146)$$

### 5.14 Contact heating

If a metal is deposited on a semiconductor with a low doping level, a Schottky contact is formed resulting in rectifying behaviour. This behaviour and the associated voltage drop is undesirable for most modern applications requiring high electrical efficiency and high modulation rates. If the semiconductor surface is doped, the difference

between the Fermi levels of the metal and the semiconductor decreases (in the absence of surface Fermi-level pinning). This decreases the barrier height and the thickness of the depletion region, resulting in a more linear IV curve. If a high enough dopant concentration is introduced into the contact region, ohmic behaviour is observed. There are three main mechanisms that control current flow across a metal/semiconductor interface [35]:

1. Thermionic emission: This mechanism is dominant in moderately doped semiconductors ( $N_d < 1 \times 10^{17} \text{ cm}^{-3}$ ), where the depletion region is wide making tunnelling a low probability event. If the barrier height is small, a large number of electrons have enough kinetic energy to surmount it and Ohmic behaviour is observed. However, if the barrier height is large and most the electrons can not surmount it, non-Ohmic behaviour is observed [35].
2. Thermionic-field emission: This mechanism dominates in moderately doped semiconductors ( $1 \times 10^{17} < N_d < 1 \times 10^{18} \text{ cm}^{-3}$ ). Thermionic emission and tunnelling both take place around the junction [35].
3. Field emission (FE): Tunnelling is the dominant carrier transfer mechanism when the semiconductor is heavily doped and the depletion region is narrow [35].

Ohmic contact resistance is a large subject. A good overview can be found in [36] and a more theoretical treatment can be found in [37] and [38]. Specific contact resistance is defined as

$$\rho = \left( \frac{\partial J}{\partial V} \right)_{V=0}^{-1} . \quad (5.147)$$

This is an experimentally determined parameter, usually determined using the transmission line method (TLM). In the TLM, a series of metal contacts (pads) are laid down separated by varying distances. The measured resistance between the pads are plotted on a graph as a function on the distance between neighbouring pads. The total resistance between the pads is the sum of two contact resistances and the bulk resistance. The line is extrapolated to the intersect and this value gives twice the contact resistance. The TLM method is, however, a thin film technique which assumes uniform current flow in the film (i.e. the impact of current spreading effects is neglected).

Tables 5.11 and 5.12 give various values of contact resistance for different material systems [36] from the literature. A good value for contact resistance is usually considered as anything around or below  $1 \times 10^{-7} \Omega \text{cm}^2$  and a poor value for contact resistance is usually considered as anything above or around  $1 \times 10^{-5} \Omega \text{cm}^2$ . Values reported in the literature for metal n-GaAs and p-GaAs interfaces are often around  $1 \times 10^{-6} \Omega \text{cm}^2$ . The contact resistance of the interface is very dependent upon the deposition technique (e.g. sputtering, evaporation or chemical vapour deposition (CVD)). Annealing has been shown to improve contact resistance, depending upon the annealing time and temperature. Because of the strong dependence of contact resistance on the processing technology used, the only way to know the contact resistance of a sample is to test a sample with the TLM. Lacking real measurement

data for the 1.3 $\mu\text{m}$  lasers, a review of the literature was conducted to find reasonable values to use (tables 5.11 and 5.12). For both n and p contacts, typical values are around  $1 \times 10^{-6} \Omega\text{cm}^2$ . Therefore, this was the value chosen as a simulation parameter.

<b>System</b>	<b>Anneal</b>	$N_a - N_a$ ( $\text{cm}^{-3}$ )	$\rho$ ( $\Omega\text{cm}^2$ )	<b>Reference</b>
Ge/Pd	Y	$1 \times 10^{18}$	$1-3 \times 10^{-6}$	[39,40]
Si/Pd	Y	$1 \times 10^{18}$	$2-6 \times 10^{-6}$	[41]
Pd/In/Pd	Y	$1 \times 10^{18}$	$1 \times 10^{-6}$	[42]
Al/Ni/Ge	Y	$1 \times 10^{18}$	$1.4 \times 10^{-6}$	[43]
Cu/Ge	Y	$1-3 \times 10^{18}$	$6.5 \times 10^{-7}$	[44]
Au/Pt/Ti	N	$1.5 \times 10^{18}$	$1.1 \times 10^{-6}$	[45]
Au/TaSiN/Au/Ge/Pt	Y	$1 \times 10^{18}$	$3.7 \times 10^{-6}$	[46]
Ag,Au/Ti	N	n/a	$2-5 \times 10^{-7}$	[47]
Au/Ge/Pd	Y	$1 \times 10^{18}$	$1 \times 10^{-6}$	[48,49]

*Table 5.11: Values for n-GaAs contacts. The values in this table were taken from [35].*

System	Anneal	$N_a-N_d$ (cm <sup>-3</sup> )	$\rho$ ( $\Omega\text{cm}^2$ )	Reference
Au/Zn/Au	Y	$3 \times 10^{17}$	$3.3 \times 10^{-6}$	[50]
Pt/Ti	N	$5 \times 10^{20}$	$8 \times 10^{-7}$	[51]
W or WSi	N	$5 \times 10^{19}$	$1 \times 10^{-6}$	[52]
Au/Zn/Au	N		$1 \times 10^{-6}$	[53]
Si/Ni/Mg/Ni	Y	$8 \times 10^{18}$	$7 \times 10^{-7}$	[54]
Pd/Sb/Mn/Pd	Y	$2.5 \times 10^{18}$	$2 \times 10^{-6}$	[55]

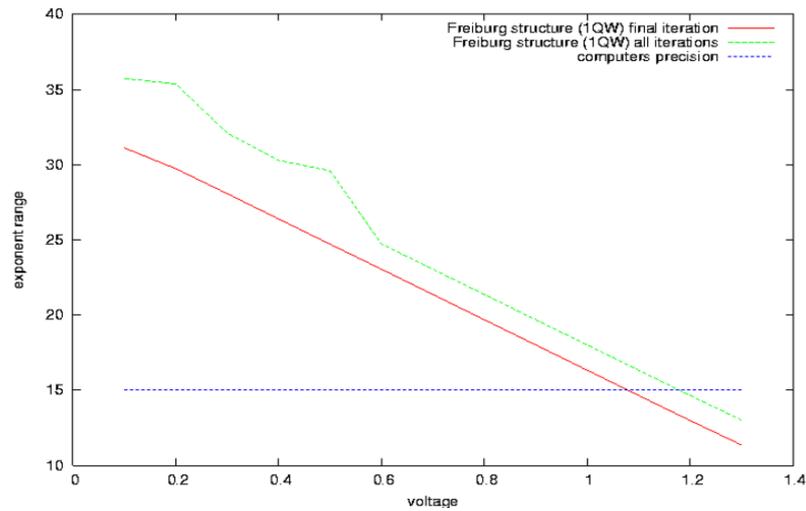
Table 5.12: Values for p-GaAs contacts. The values in this table were taken from [35]

The heat generated from the contact resistance was included in the simulator at the metal/semiconductor interface as a heat source. The voltage drop across the contact is added to the total applied voltage.

### 5.15 Stability of the electrical model

At low currents (i.e. below threshold), spikes were observed in the current density profile. Upon approaching threshold, the simulator performed well. Below threshold, numbers with a wide range of magnitudes ( $1 \times 10^{-25} - 1 \times 10^{10}$ ) were found in the Jacobian (figure 5.8). The mantissa (floating point precision) of a double precision number in a PC is only 53 bits (15 digits). Thus, the problem exceeded the computers floating point precision. This section describes the work done to correct this problem, including the implementation of back tracking, clamping and matrix normalisation. Although these techniques provided increased stability, it was found that the solution time was slower. Therefore, these techniques should only be used when convergence

is hard to achieve.



*Figure 5.8: A plot of the range of exponents in the Jacobian when the simulator is run on a laser structure developed by the Fraunhofer Institute for Applied Solid State Physics [56]. It can be seen from this graph that the numbers within the Jacobian only fall within the precision of the computer for voltages above 1.1 volts (threshold).*

### 5.15.1 Matrix normalisation

In an effort to reduce the matrix bandwidth, matrix normalisation was performed. Each row was divided by the diagonal element, resulting in a diagonal of 1. This was found to reduce the matrix bandwidth by up to two orders of magnitude.

### 5.15.2 Iterative improvement of the solution

Direct methods of solving linear equations often result in large round-off errors. In

order to improve the accuracy of the LU sparse decomposition algorithm, an iterative improvement of the solution was performed [57]. Firstly, the set of linear equations

$$\mathbf{Ax} = \mathbf{b} \quad (5.148)$$

is rearranged to give

$$\mathbf{Ax} - \mathbf{b} = \mathbf{0} \quad (5.149)$$

$\mathbf{A}\delta\mathbf{x}$  is then added to each side

$$\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) - \mathbf{b} = \mathbf{A}\delta\mathbf{x} \quad (5.150)$$

where  $\mathbf{x} + \delta\mathbf{x}$  is the real solution vector to the set of linear equations with no computational error plus some error  $\delta\mathbf{x}$ . The solution given by sparse LU decomposition is  $\mathbf{x} + \delta\mathbf{x}$ . After the initial solution step, all of the parameters on the left hand side of equation 5.150 are known. Equation 5.150 can then be solved for  $\delta\mathbf{x}$ . This can be subtracted from the first approximate solution to obtain a a better solution with less rounding error. The process can be repeated until a desired level of accuracy is obtained. The procedure can gain a significant speed up if one notes that LU decomposition has already been preformed on the Jacobian  $\mathbf{A}$ . Thus the program can skip straight to the solution routine to obtain  $\delta\mathbf{x}$ . Under some circumstances, this method is essential to achieve convergence. However it requires more computational steps. Thus, it was left as a option only to be used when convergence is hard to obtain.

### 5.15.3 Clamping and back tracking algorithms for fast global convergence

The general idea of back-tracking is to note that Newton's method always produces a

step in the right direction (i.e. towards the solution). However, the step is sometimes too big and the solution is missed. This can lead to slow convergence or even no convergence at all. A back-tracking algorithm, evaluates the Newton step. If it is acceptable, no change is made to the step vector. However, if the step is too big and *increases* the error, the step vector is multiplied by a constant in order to make the global error decrease. The thermal solver was generally found to be stable and not require clamping or back-tracking. However, the Newton solver of the main laser diode simulation software was shown to sometimes benefit from clamping and back-tracking to force global convergence.

A back-tracking algorithm from [58] was used and will now be described. The Newton scheme is defined as

$$\delta x = -\mathbf{J}^{-1} \cdot \mathbf{F} \quad (5.151)$$

where  $\mathbf{F}(\mathbf{x})$  is the error vector,  $\delta x$  is the update and  $\mathbf{J}$  is the Jacobian, which is defined as

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial x} \quad (5.152)$$

An error function can be written as

$$f(x) = \frac{1}{2} \mathbf{F} \cdot \mathbf{F} \quad (5.153)$$

Differentiating  $f(x)$  results in

$$f'(x) = \frac{1}{2} \frac{\partial \mathbf{F}}{\partial x} \cdot \mathbf{F} + \frac{1}{2} \frac{\partial \mathbf{F}}{\partial x} \cdot \mathbf{F} \quad (5.154)$$

and adding the result gives

$$\frac{\partial f}{\partial x} = \mathbf{J} \cdot \mathbf{F} \quad . \quad (5.155)$$

Equation 5.156 proves that the Newton step is always in the correct direction to minimise the error

$$\frac{\partial f}{\partial x} \delta x = (\mathbf{J} \cdot \mathbf{F}) \cdot (-\mathbf{J}^{-1} \cdot \mathbf{F}) = -\mathbf{F} \cdot \mathbf{F} < 0 \quad . \quad (5.156)$$

For a Newton step to be accepted, the error  $f$  must decrease. However, the step may be too big and need scaling in order to produce an optimum step. If the full Newton step is defined as  $\mathbf{p}$  and  $\lambda$  is defined as a scaling factor, then new state variables can be defined as

$$\mathbf{x}_{new} = \mathbf{x}_{old} + \lambda \mathbf{p} : \quad 0 \leq \lambda < 1 \quad . \quad (5.157)$$

To decide whether or not to accept a step, the error  $f$  must decrease and the decrease must not be too slow. To avoid this, the criteria

$$f(\mathbf{x}_{new}) \leq f(\mathbf{x}_{old}) + \alpha \nabla f \cdot (\mathbf{x}_{new} - \mathbf{x}_{old}) \quad . \quad (5.158)$$

is set. This forces the error function  $f(\mathbf{x}_{new})$  to be at least equal to the old error function  $f(\mathbf{x}_{old})$  minus some fraction of the initial decrease of the error -  $\alpha$  is usually set to  $1 \times 10^{-4}$  [58]. The error at the new point in the multi dimensional vector space is defined as

$$g(\lambda) = f(\mathbf{x}_{old} + \lambda \mathbf{p}) \quad , \quad (5.159)$$

and its differential with respect to  $\lambda$  is defined as

$$g'(\lambda) = \nabla f \cdot \mathbf{p} \quad . \quad (5.160)$$

The first step is to check if  $\lambda=1$  is an acceptable option using equation 5.158. If this is not acceptable, then  $g(1)$  is now known as well. The error function can  $g(\lambda)$  be

modelled as a quadratic equation

$$g(\lambda) \approx [g(1) - g(0) - g'(0)]\lambda^2 + g'(0)\lambda + g(0) \quad . \quad (5.161)$$

The minimum in the error can be found by taking the derivative with respect to  $\lambda$  and solving for lambda. This gives

$$\lambda = \frac{g'(0)}{2[g(1) - g(0) - g'(0)]} \quad . \quad (5.162)$$

The scaling factor,  $\lambda$  is not allowed to be smaller than 0.1, so that convergence does not grind to a halt. If the value for  $\lambda$  is not accepted (using equation 5.158), then for the next iteration  $\lambda$  is modelled as a third-order polynomial

$$g(\lambda) = a\lambda^3 + b\lambda^2 + g'(0)\lambda + g(0) \quad . \quad (5.163)$$

Using the same approach, the minimum of this cubic polynomial is found from

$$\lambda = \frac{-b + \sqrt{(b^2 - 3ag'(0))}}{3a} \quad . \quad (5.164)$$

To protect the Newton solver from very large steps, the correction to the state variables is first coarsely clamped before the back-tracking algorithm is called. The following clamping scheme is used

$$\mathbf{x}_{clamp} = \frac{1}{1.0 + \left| \frac{\mathbf{x}}{clamp} \right|} \quad . \quad (5.165)$$

This allows the back-tracking algorithm to adjust the step size within the limit set by the clamp. The use of back-tracking reduces the total number of iterative steps needed to solve the semiconductor equations and also decreases the total error. However the disadvantage is that every time  $g(x)$  is evaluated, the error function for every equation in the matrix must be re-evaluated. This is a slow process, so the back

tracking algorithm is only triggered if the error at the end of a solution without back tracking is judged to be unacceptable.

#### **5.15.4 Bi-CGSTAB**

An iterative sparse matrix solver was tested as a possible replacement for the sparse LU decomposition [60] combined with the direct sparse solver used to solve the electrical problem. The stabilised bi-conjugate method (Bi-CGSTAB) with LU preconditioning [59] has a good reputation for being a stable and fast matrix solver. However, it proved to require very accurate LU decomposition in the preconditioning step to guarantee convergence of the Newton solver. Bi-CGSTAB was thus found to be impractical.

#### **5.16 Summary**

In this chapter, an isothermal device simulator has been extended to include thermal effects. The lattice heat equation has been discretised and solved with a coupled iterative solution of the electrical and thermal problems. A database capable of storing and using the temperature dependent material parameters has been written. Band gap energy, thermal conductivity, electron affinity, recombination constants, gain, spontaneous emission and mobility all vary as a function of temperature within the model. The need for a unified electrical and thermal mesh has been demonstrated. A model which takes into account the energy lost during the capture process has been derived.

The thermodynamic treatments and hydrodynamic models were reviewed and their relation to the drift diffusion model has been shown. The current equations have been discretised and implemented into the device simulator using a Scharfetter-Gummel discretisation scheme. Finally, the stability of the electrical model has been examined. Matrix normalisation, iterative improvement of the solution, clamping, back-tracking algorithms and the Bi-CGSTAB matrix solver have been evaluated as optimisation techniques.

The result is a simulation tool capable of full electro-thermal-optical simulation. In the next chapter, the model will be compared with experimental measurements reported in Chapter 4. Finally, the resulting model will be used to optimise the thermal environment of the 1.3 $\mu\text{m}$  edge-emitting lasers investigated in Chapters 3-4, with the aim of designing devices with less self-heating and therefore capable of longer lifetimes.

## **5.17 References**

- [1] M. Knaipp, "Modellierung von Temperatureinflüssen in Halbleiterbauelementen", Technischen Universität Wien, September 1998 English: "Moderling of temperature effects in semiconductor devices", Technical University of Viena September 1998
- [2] S. Sujecki, L. Borrueal, J. Wykes, P. Moreno, B. Sumpf, P. Sewell, H. Wenzel, T.M. Benson, G. Erbert, I. Esquivias, E.C. Larkins, "Nonlinear properties of tapered laser cavities" IEEE J. Sel. Top. in Quant. Electron., May-June 2003, 9,3 pp. 823-834
- [3] M. Sotoodeh, A. H. Khalid, and A. A. Rezazadeh "Empirical low-field mobility

model for III-V compounds applicable in device simulation codes", Journal of Applied Physics, 87,6, pp. 2890-2900, 2000

[4] J. Piprek, "Semiconductor Optoelectronic Devices: Introduction to Physics and Simulation", Academic Press, ISBN 0125571909, 2000

[5] S. Selberherr, "Analysis and Simulation of Semiconductor Devices", Springer-Verlag, ISBN 3-211-40537-2, p.43,1984

[6] D. R. Lide, "CRC Hand book of Chemistry and physics 82nd edition", CRC press ,ISBN: 0849304822, 2001-2002

[7] D. R. Lide, "CRC Hand book of Chemistry and physics 82nd edition", CRC press ,ISBN: 0849304822, 2001-2002

[8] S. Selberherr "Analysis and Simulation of Semiconductor Devices", Springer-Verlag, ISBN 3-211-40537-2, p.42, 1984

[9] I. Vurgaftmana, J. R. Meyer and L. R. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys" Journal of Applied Physics, 89, 11, 2001

[10] S. Adachi, "GaAs, AlAs, and  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  Material parameters for use in research and device applications", J. Appl. Phys., 58, 3, pp. R1-R29, 1985

[11] D. N. Nichols, D. S. Rimai and R. J. Sladek, "Elastic anharmonicity of InP: Its relationship to the high pressure transition", Solid State Commun., 36, 8, pp. 667-669, 1980

[12] S. Selberherr "Analysis and Simulation of Semiconductor Devices", Springer-Verlag, ISBN 3-211-40537-2, 1984

[13] H.C. Casey and M.B. Panish , "Heterostructure Lasers Part A: Fundermental Principles", Academic Press ,ISBN 0-12-163191-X ,p. 160,1978

- [14] J. Piprek, "Semiconductor Optoelectronic Devices Introduction to Physics and Simulation", Academic Press, ISBN 0387226591, p. 71, 2005
- [15] S. Chinn, P. Zory and A.R. Reisinger, "A Model for GRIN-SCH-SQW Diode Lasers" IEEE J. Quant. Electron., 24, 11, pp. 2191, 1988
- [16] S. Chinn, P. Zory, A.R. Reisinger, "GRIN-SCH-SQW Diode Lasers", Quant. Electron., 24, 11, pp. 2191, 1988
- [17] D. R. Lide, "CRC Hand book of Chemistry and physics 82nd edition", CRC press ,ISBN: 0849304822, 2001-2002
- [18] G. K. Wachutka, "Rigorous Thermodynamic Treatment of Heat Generation and Conduction in Semiconductor Device Modeling" IEEE Transactions on computeraided design, 9, 11, pp. 1141-1149, 1990
- [19] G. Wachutka, "Unified framework for thermal, electrical, magnetic, and optical semiconductor device modeling.", COMPEL, 10, pp. 311-321, 2004
- [20] S. Rudin, G. Wachutka and H Baltes, "Thermal Effects in Magnetic Microsensor Modeling", Sensors and Actuators A. pp. 731-735, 1991
- [21] R. Stratton, "Diffusion of Hot and Cold Electrons in Semiconductor Barriers ", Phys. Rev. ,126,6, pp.2002 - 2014, 1962
- [22] E.M. Azoff, "Generalized Energy-Momentum conservation equations in the relaxation time approximation", Solid-State Electron., 30,9, pp. 913-917,1987
- [23] Minimos-NT home page, <http://www.iue.tuwien.ac.at/index.php?id=2>, Accessed 14 Feb 08
- [24] T. Grasser, Ting-Wei Tang, H.Kosina and S.Selberherr, "A Review of Hydrodynamic and Energy-Transport Models for Semiconductor Device Simulation",

Proceedings of the IEEE, 91, 2, p. 251, 2003

[25] K. Kells "General Electrical Semiconductor Device Simulation", Hartung-Gorre Verlag Konstanz - Series in Microelectronics, ISBN 3-89191-787-2,1994

[26] K. Bløtekjær "High-frequency conductivity carrier waves and acoustic amplification in drifted semiconductor plasmas", Ericson Technics, 22, 2, p. 125

[27] R. Kosik, "Numerical Challenges on the Road to NanoTCAD", Dissertation, Wien <http://www.iue.tuwien.ac.at/phd/kosik/thesis.html>, accessed 14 Feb 08

[28] S. Selberherr, V. Palankovski and R. Quay, "Analysis and Simulation of Heterostructure Devices - Computational Microelectronics", Springer-Verlag, Wien, ISBN 3-211-40537-2 ,2004

[29] D.L. Scharfetter, H.K. Gummel, "Large-signal analysis of a silicon Read diode oscillator", IEEE Transactions on Electron. Devices, 16, 1, pp. 64-77,1969

[30] P. S. Zory "Quantum Well Lasers", Academic Press,ISBN 0127818901, 1993

[31] S. Selberherr, V. Palankovski and R. Quay, "Analysis and Simulation of Heterostructure Devices - Computational Microelectronics", Springer-Verlag, Wien, ISBN 3-211-40537-2 ,2004

[32] J. E. Carroll, J. Whiteaway, D. Plumb "Distributed feedback semiconductor lasers" Published by IET, ISBN 0852969171, 1998

[33] M. Gault, P. Mawby, A.R. Adams, M. Towers, "Two-Dimensional simulation of Constricted-Mesa InGaAsP/InP Buried-Heterostructure Lasers", IEEE J. Quant. Electron. ,30, 8, pp. 1691 , 1994

[34] H.C. Casey and M.B. Panish , "Heterostructure Lasers Part A: Fundamental Principles", Academic Press ,ISBN 0-12-163191-X ,p. 174, 1978

- [35] T.C. Shen G.B. Gao and H. Morko, "Recent developments in ohmic contacts for III-V compound semiconductors", *J. Vac. Sci. Technol. B*, 10,5, pp. 2113, 1992
- [36] A. G. Baca, F. Ren, J. C. Zolper, R. D. Briggs and S. J. Pearton , "A survey of ohmic contacts to III-V compound semiconductors" ,*Thin Solid Films*, 308-309 , pp. 599-606, 1997
- [37] V.L. Rideout, "A review of the theory and technology for ohmic contacts to group III-V compound semiconductors", *Solid State Electron*, 18, 541-550, 1975
- [38] A. Piotrowska, A. Guivarch and G. Pelous, *Solid State Electron.* , "Ohmic contacts to III-V compound semiconductors: A review of fabrication techniques", 26, pp. 179-197, 1983
- [39] E. D. Marshall, W. X. Chen, C. S. Wu, S. S. Lau, and T. F. Kuech, "Non-alloyed ohmic contact to n-GaAs by solid phase epitaxy " *Appl. Phys. Lett.*, 47, pp. 298-300, 1985
- [40] E. D. Marshall, B. Zhang, L. C. Wang, P. F. Jiao, W. X. Chen, T. Sawada, S. S. Lau, K. L. Kavanagh and T. F. Kuech "Nonalloyed ohmic contacts to n-GaAs by solid-phase epitaxy of Ge" , *J. Appl. Phys.*, 62, 942, p. 1987
- [41] L.C. Wang, et al. *J. Mater. Res.*, 3, p. 922, 1988
- [42] L. C. Wang, X. Z. Wang, S. S. Lau, T. Sands W. K. Chan, and T. F. Kuech, "Stable and shallow PdIn ohmic contacts to n-GaAs", *Appl. Phys. Lett.*, 56, p.2129, 1990
- [43] R. Zuleeg, et al. , *IEEE Electron Dev. Lett.*, 7, p.603, 1996
- [44] M.O. Aboelfotoh, C.L. Lin, J.M. Woodall, "Novel low-resistance ohmic contact to n-type GaAs using Cu<sub>3</sub>Ge", *Appl. Phys. Lett.*, 65, p.3245, 1994

- [45] F. Ren, A.Y. Cho, D.L. Sivco, S.J. Pearton, C.R. Abernathy, "Use of Sn-doped GaAs for non-alloyed ohmic contacts to HEMTs", *Electron. Lett.*, 30,11, pp. 912-914, 1994
- [46] J.S. Chen, E. Kolawa, M.-A. Nicolet, R.P. Ruiz, "Current generated periodic space charge in semiconductor diodes with non-Ohmic injection", *J. Appl. Phys.*, 75, p.373, 1994
- [47] M.P. Patkar, T.P. Chin, J.M. Woodall, M.S. Lundstrom, M.R. Melloch, "Very low resistance nonalloyed ohmic contacts using low-temperature molecular beam epitaxy of GaAs", *Appl. Phys. Lett.*, 66, p.1412, 1995
- [48] L.C. Wang, P.H. Hao, B.J. Wu, "Low-temperature-processed (150-175 °C) Ge/Pd-based Ohmic contacts ( $c \sim 1 \times 10^{-6} \text{ cm}^2$ ) to n-GaAs", *Appl. Phys. Lett.*, 67, p. 509, 1995
- [49] L.C. Wang, P.H. Hao, J.Y. Cheng, F. Deng and S.S. Lau, "Ohmic contact formation mechanism of the Au/Ge/Pd/n-GaAs system formed below 200 °C", *J. Appl. Phys.*, 79, p. 4216, 1996
- [50] Y. Lu, T. S. Kalkur, and C. A. Paz de Araujo, "Rapid Thermal Alloyed Ohmic Contacts to p-Type GaAs", *J. Electrochem. Soc.*, 136, pp. 3123-3129, 1989
- [51] A. Katz, C.R. Abernathy, S.J. Pearton, "Pt/Ti ohmic contacts to ultrahigh carbon-doped p-GaAs formed by rapid thermal processing", *Appl. Phys. Lett.*, 56, p. 1028, 1990
- [52] M. L. Lovejoy, J. C. Zolper, M. E. Sherwin, A. G. Baca, R. J. Shul, D. J. Rieger and J. F. Klem, "Non-alloyed, refractory metal contact optimisation with shallow implantations of Zn and Mg", *Thin Solid Films*, 253 pp. 496-500, 1994

- [53] T.S. Kalkur, Y.C. Lu, C.A. Araujo, "Non-alloyed ohmic contacts on rapid thermally Zn diffused GaAs" *Solid State Electron.*, 32, pp. 281-285, 1989
- [54] C.C. Han, X.Z. Wang, L.C. Wang, E.D. Marshall, S.S. Lau, S.A. Schwarz, C.J. Palmstrom, J.P. Harbison, L.T. Florez, R.M. Potemski, M.A. Tischler, T.F. Kuech, "Nonspiking ohmic contact to p-GaAs by solid-phase regrowth", *J. Appl. Phys.*, 68 p. 5714, 1990
- [55] C.C. Han, X.Z. Wang, S.S. Lau, R.M. Potemski, M.A. Tischler, T.F. Kuech, "Thermally stable and nonspiking Pd/Sb(Mn) ohmic contact to p-GaAs", *Appl. Phys. Lett.*, 58 p. 1617, 1991
- [56] S. Burkner, J. D. Ralston, S. Weisser, J. Rosenzweig, E. C. Larkins, R. E. Sah, and J. Fleissner, "Wavelength Tuning of High-speed InGaAs-GaAs-AlGaAs Pseudomorphic MQW Lasers Via Impurity-Free Interdiffusion" , *Photonics Technology Letters*, 7,9, p. 941,1995
- [57] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery "Numerical recipes in C: the art of scientific computing", Cambridge University Press, ISBN 0521437202, p. 56, 2002
- [58] W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery "Numerical recipes in C: the art of scientific computing", Cambridge University Press, ISBN 0521437202,2002
- [59] NAG Fortran Library Manual Mark 21, Numerical Algorithms Group, Section f11, Routine: f11dcf
- [60] NAG Fortran Library Manual Mark 21, Numerical Algorithms Group, Section f04, Routine: f04axf