# Investigations into thermionic cooling for domestic refrigeration

Benjamin C.C. Lough
University of Wollongong

# INVESTIGATIONS INTO THERMIONIC COOLING FOR DOMESTIC REFRIGERATION

by

Benjamin C. C. Lough

A thesis submitted in fulfilment of the
requirements for the Doctor of Philosophy
degree in the School of Engineering Physics at the
University of Wollongong

2004

Thesis supervisors: Associate Professor Roger Lewis and
Professor Chao Zhang

CERTIFICATION

I, Benjamin C. C. Lough, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Engineering Physics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

_____

Benjamin C. C. Lough

July 5, 2004

**ACKNOWLEDGEMENTS**

First and foremost I must thank my supervisors, Associate Professor Roger Lewis and Professor Chao Zhang for their continued support and advice throughout the course of this work. I thank Professor Lewis for his direction for the work and for his guidance with both experimental and theoretical aspects of the work carried out. I thank Professor Zhang for the long discussions regarding the theory behind thermionic cooling and physics in general. Both offered a lot of their time to me and without their availability I would not have been able to complete the work that I did. I thank them both for always making me feel welcome in their offices. I also thank them for giving me the opportunity to present work at an International Conference in France.

I thank Sueping Lee for the experimental work she carried out on the first two generations of devices that were studied in this work and the data she made available for analysis. I also thank her for her encouragement and our friendship that has developed during this time.

I would like to extend my thanks to Professor Chennupati Jagadish and Dr Hoe Tan at the Australian National University for help on the design and growth of the first two generations of devices studied in this project.

Thankyou to Dr Scott Butcher at Macquarie University for lending our team a wire bonder for use with these devices.

I would like to thank Mr Xunzhong Shang from the Institute of Physics, Chinese Academy of Sciences, and Associate Professor Qian Wang from Xinjiang University for their discussions and suggestions for material selection and fabrication techniques for the fourth-generation devices proposed in this work.

Thanks to Dr Iwan Cornelius at the University of Wollongong for the discussions we had over coffee about everything and nothing at the same time.

I thank Mr Peter Anthony for the use of tools and components from his workshop and Mr John Brady for advice on daisy-chaining multiple digital spectrometers.

Also I thank everyone in the School of Engineering Physics, University of Wollongong, who at one time or another offered their time and help to me.

Finally I would like to thank my parents and family for their encouragement and support throughout this work. I thank Allysia for putting up with me whilst writing this and making the long process more enjoyable.

I would like to acknowledge the financial support of the Australian Research Council, The University of Wollongong and Email Ltd.

## ABSTRACT

Thermionic refrigeration using semiconductor heterostructures is examined theoretically and experimentally. A theory of single-barrier devices is first developed where two classes of single-barrier devices are defined and compared. So-called class 1 devices are found to always perform better. A theory of multiple-barrier devices based on class 1 barriers is then developed using a numerical solution.

Experimentally, three generations of 10-barrier devices based on $Al_xGa_{1-x}As$-GaAs heterostructures were made and electrically characterised. This material is by no means ideal (as will be discussed) but was used due to availability and because, at the commencement of this work, had never been used for this purpose before. Thermal measurements were made to determine if any cooling occurred at room temperature. No cooling was observed but the electrical characteristics allowed for examination of the models developed. It was found that the earlier models used did not accurately model the $I$-$V$ characteristics of the devices. This was attributed to the fact that the initial models did not take space-charge into account.

A more robust numerical model is developed in which the $I$-$V$ characteristics of devices are predicted much more accurately. This model is then used to design new generations of devices. The work concludes by recommending a next generation design in which substantially more cooling is expected compared to the samples examined here. The probability of cooling being observed in the future is therefore increased.

The type of devices described here will always be hindered because of heat conduction. Other methods incorporating thermionic emission, such as an opto-thermionic system in which removed heat is given off as light, may ultimately prove to be the best solution. This aside, it is hoped that the work presented here will contribute to the understanding of the field.

## TABLE OF CONTENTS

## LIST OF TABLES

Table Page

## LIST OF FIGURES

## LIST OF PUBLICATIONS

### Refereed Publications

1. B. C. Lough, S. P. Lee, R. A. Lewis and C. Zhang, "Electronic thermal transport and thermionic cooling in semiconductor multi-quantum-well structures," *Computer Physics Communications*, vol. 142, pp. 274 – 280, 2001.
2. B. C. Lough, S. P. Lee, R. A. Lewis and C. Zhang, "Numerical calculation of thermionic cooling efficiency in a double-barrier semiconductor heterostructure," *Physica E: Low-Dimensional Systems and Nanostructures*, vol. 11, pp. 287 – 291, 2001
3. B. C. Lough, S. P. Lee, Z. Dou, R. A. Lewis, C. Zhang, "Multilayer thermionic cooling in semiconductor heterostructures," Australian Institute of Physics Congress 2002 - Physics and Industry working together, 2002, *Proceedings of the Australian Institute of Physics 15th Biennial Congress 2002 - Physics and Industry working together*, Australian Institute of Physics, Sydney, Australia, pp.364–366.
4. C. Zhang and B. Lough, "Dynamical conductivity of a two-dimensional electron gas under an intense terahertz radiation," *Physica E*, vol. 17, pp. 289 – 291, 2003
5. B. C. C. Lough, S. P. Lee, Z. Dou, R. A. Lewis and C. Zhang, "Investigation into space charge effects in $I$-$V$ characteristics of multi-layer semiconductor thermionic devices," *Physica E*, vol. 17, pp. 651 – 653, 2003

### Conferences

1. S. P. Lee, B. C. Lough, R. A. Lewis and C. Zhang, "Thermionic cooling of optoelectronics and microelectronic devices," *2000 Conference on Optoelectronics and*

*Microelectronics Materials and Devices*, 6–8 December, 2000, Melbourne, Australia.

2. S. P. Lee, B. C. Lough, R. A. Lewis and C. Zhang, "Multilayer Thermionic Cooling in Semiconductor Heterostructures," *$26^{th}$ ANZIP Condensed Matter Meeting*, 29 January – 1 February, 2002, Wagga Wagga, Australia.

3. R. A. Lewis and B. C. .C Lough, "Spectroscopic determination of valence band parameters of InP," *Proc. 27th Annual A&NZIP Cond. Matt. & Mat. Meeting*, 4 February – 7 February, 2003, Wagga Wagga, Australia.

4. B. C. C. Lough, S. P. Lee, Z. Dou, R. A. Lewis and C. Zhang, "Thermal and electrical characteristics of a multilayer thermionic device," *Proc. of Optoelectronic and Microelectronic Materials and Devices*, 11 December – 13 December, 2003, Sydney, Australia, pp.511–514

# CHAPTER 1

# INTRODUCTION

Table 1.1: Value of Fundamental Constants

| **Constant** | **Description** | **Value** | **Unit** |
|---|---|---|---|
| $q$ | Charge of Electron | $1.6022 \times 10^{-19}$ | C |
| $k_B$ | Boltzmann's Constant | $1.3807 \times 10^{-23}$ | J/K |
| $\epsilon_0$ | Permittivity of Free Space | $8.85 \times 10^{-12}$ | $C^2/N{\cdot}m^2$ |
| $m_0$ | Mass of Free Electron | $9.1094 \times 10^{-31}$ | kg |

## 1.1 Thermionic Emission

Thermionic Emission is by no means a new concept. It was first discovered by Thomas Edison in 1883 when he was working with filaments for the electric light. Thermionic Emission refers to electrons or ions being ejected from a heated surface. As these *therm-ions* leave the surface, they take heat with them in the form of kinetic energy. In doing so, they cool the surface they are leaving. This forms the basis of using thermionic emission to achieve cooling.

The simplest thermionic device comprises of two metal plates, separated by a distance, that are held at different temperatures. Both plates will eject thermions, but there will be a net flow of electrons from the hot to cool plate, as more thermions leave the plate held at the elevated temperature. Alternatively, a net flow of electrons can be induced between two plates held at the same temperature if a potential is applied across the two plates. This potential will aid the flow of electrons leaving one plate and will retard the electrons leaving the other. By doing this a net electrical, and an accompanying energy,

flow will occur between the two plates. The plate from which the net current flows will become cooler, while the plate that the net current flows to will become hotter. The aim of using thermionic emission for cooling is to increase the efficiency and power output of current solid-state devices. The efficiency of the device is the ratio of the cooling power achievable (in the form of the net heat flow from the cooling surface) to the amount of input power required to keep a net current flowing from the cooling surface.

### 1.2 Motivation for this work

The successful development of semiconductor-based thermionic coolers will see a great number of applications for the new technology. With the current concern about global warming due to green house gases a replacement for CFCs is required. Some of these replacement gases are either powerful green house gases themselves, poisonous or dangerous. It is feasible that gas-based refrigeration systems could be replaced altogether by a semiconductor-based electronic system such as thermoelectric or thermionic devices. Not only would they replace the need for gases but would also have a number of other very desirable properties. Having no moving parts, these system would by silent and require little maintenance due to wear. It is also foreseeable that these devices may one day be more efficient than gas-based systems.

With advances in semiconductor laser sources for the telecommunications industry, highly efficient devices that have low heat dissipation are now common. Temperature dependent parameters such as wavelength drift [1] can therefore be stabilised using little cooling [2]. Conventional themoelectric coolers have been used to achieve this cooling in the past but are not cost effective due to packaging costs. Semiconductor coolers made from the same material that can be grown directly onto the devices may offer a cost-effective solution.

With decreased size and increased speed of microprocessors, heat management has also become an important issue in the performance of these devices. By using the same fabrication techniques as for the devices themselves it is believed integrated coolers may offer

a solution to heat management.

Another promising application for thermionic devices is in the area of power generation. Cooling in thermionic devices is achieved by applying a bias across the device causing a net heat current to flow from one electrode to the other. By applying a temperature gradient across the device electrons are thermally excited over the potential barrier and a net electric current flows [3]. It is possible that such devices could harness waste heat of conventional power generators to increase their overall efficiency.

### 1.3 Thermoelectric and Thermionic Cooling

Thermoelectric and thermionic cooling are very similar. Both achieve cooling by a net electric current causing energy to be transported by electrons. Thermoelectric cooling results from the bulk properties of materials [4] as electrons move diffusively through the material. Thermionic cooling, on the other hand, assumes that electrons move ballistically across a potential barrier. By doing so, the electrons carry all of their kinetic (thermal) energy from one electrode to the next and thus theoretically have a higher efficiency [3].

The efficiency for thermoelectric coolers is determined by the thermoelectric figure of merit, $Z$, given by [5]:

$$Z(T) = \frac{S^2}{\rho K} \tag{1.1}$$

where $\rho$ is the electrical resistivity, $S$ is the Seebeck coefficient and $K$ is the thermal conductivity. Generally these parameters are temperature dependent. Thermoelectric devices lose efficiency because $Z$ varies over the length of the device [6]. It has been proposed that this efficiency can be increased by using segmented devices [7, 8]. Work carried out by Hicks et al. [9] showed that the use of multiple-quantum-well structures increased the thermoelectric figure of merit of certain materials. Mahan et al. [3] suggested that devices such as these may have increased efficiency because the barrier periods used are so short that electron motion is probably ballistic i.e. thermionic rather than thermoelectric effects

are responsible for cooling.

### 1.3.1 Semiconductor Thermionic Refrigeration

The main reason that thermionic devices have not already seen practical use since the discovery of thermionic emission is because of a factor pointed out by Mahan [10] in 1994. All thermionic devices are based on Richardson's equation which describes the current per unit area emitted by a metal surface with work function, $\phi$, at temperature, $T$:

$$J(\phi, T) = AT^2 \exp\left(-\frac{q\phi}{k_B T}\right) \tag{1.2}$$

where $A = \dfrac{qm^* k_B^2}{2\pi^2 \hbar^3} \simeq -120$ A/cm$^2$ is Richardson's constant.

Domestic refrigerators require a cold side at $\sim 250$ K. To achieve this the cathode of the device needs to emit around 1 A·cm$^{-2}$, which requires that the work function be less than $0.34$ eV [10]. This value is much lower than for any known metal-vacuum interface. In this paper, Mahan stated his belief that any system based on thermionic emission using a vacuum diode would never be sufficiently efficient at room temperature. He showed that the work function between a metal-vacuum interface can never be reduced below a minimum value of 0.71 eV at room temperature because of space-charge accumulation.

Shakouri and Bowers [11, 12] then proposed using semiconductor heterostructures to achieve thermionic cooling. They noted that problems with space charge can be controlled by modulation doping or bandgap engineering. They also noted, however, that using a semiconductor-semiconductor system causes problems due to the much higher thermal conductivity of the semiconductor material as opposed to the vacuum. Efficiency is reduced because of conduction heat flow. They suggested that cascading the devices or distributing the temperature gradient over larger distances might increase the efficiency. They also discussed material selection and maximum cooling power achievable for single-barrier devices.

Soon after these papers, Mahan [3, 13] showed that any single-barrier devices would have

to be thin to ensure ballistic transport and so the temperature gradient across the barrier would have to be small. He proposed using multi-barrier devices in order to achieve macroscopic cooling and provided a complete theory for the system.

Experimentally, Shakouri et al. [14] demonstrated cooling of around a degree in single-barrier InGaAsP, 1 $\mu$m thick, devices. LaBounty et al. [1] demonstrated around the same amount of cooling in a 25-barrier structure made of the same material, this time with barrier widths only 30 nm. They also showed that contact resistance and resistance in wire bond adversely effects results. Ulrich [15] also showed that even small resistance in the contacts will severely reduce cooling the efficiency of the devices.

### 1.3.2 Thermal Transport in Superlattices

In addition to the argument that Mahan proposed in relation to multilayer devices performing better than single-barrier devices, there has been substantial work carried out on measuring the thermal conductivity of superlattices as compared to bulk materials. Swartz [16] reported measurements of solid-solid thermal boundary resistance. He stated that at low temperature the boundary resistance is said to be in agreement with the acoustic mismatch model. At higher temperatures the influence of diffuse scattering at the interface is found to decreases less rapidly with increasing temperature than predicted by the theory. Phonon scattering in thin ($\sim$30 Å) disordered layers near the interface was shown to be a possible explanation. Since then, work has been carried out in measuring thermal conductivity in superlattices [17] and a large amount of work has also been carried out in trying to develop a model that correctly explains these results [18, 19, 20].

Huang et al. [21] showed that having superlattice periods of non-equal length and following a Fibonacci sequence reduces optical-phonon transport due to localisation of phonons. Any structure with non-equal period lengths may also have reduced thermal conductivity because of the inability of phonons to propagate through the system. Just as band engineering is used to improve electrical transport in devices, Chen et al. [22] discussed the

possibility of using phonon engineering in order to achieve low thermal conductivities but also stated this area is still in its' infancy and much more work is needed.

### 1.3.3 Novel Device Designs

Novel designs have been presented by a number of authors. Zhou et al. [23] showed that, theoretically, a varying current density allows a greater temperature gradient to be obtained across a thermionic device. Such varying current density can be obtained by using a pyramid-like structure with the cross-section of layers being increased or decreased.

The most important problem with the devices described so far is the decrease in efficiency and cooling power due to thermal conduction. As heat is removed from the cold side of the device, the temperature increases through the device. This will always remain a problem as long as heat removed remains in the device. Shakouri and Bowers [11] suggested using optical emission to remove the extracted heat from the device. Mal'shukov and Chao [24] presented work on a hybrid system combining the principles of laser cooling and thermionic cooling in order to remove extracted heat by means of Auger recombination and calculated cooling rates of at least several W/cm$^2$ using a single junction.

Hishinuma et al. [25] took another approach to hybrid systems. They suggested using a combined tunnelling and thermionic emission system to select only high energy electrons for contribution to current. They calculated that relatively high work function materials ($\sim$ 1 eV) could be used to achieve cooling power of up to 100 W/cm$^2$. The same team also proposed using a semiconductor heterojunction-vacuum interface [26] for thermionic emission. They stated that the vacuum-semiconductor potential is significantly reduced due to the Schottky effect and because of penetration of the electric field into the semiconductor. Because of this potential lowering there is no need for the low work-function materials considered before and high emission currents become possible under an applied electric field. They calculate that cooling power of 10-100 W/cm$^2$ is expected at room temperature. Xuan [27] presented work on a combined thermionic-thermoelectric

cooler and showed that it performs better and more efficiently than either thermoelectric or thermionic devices alone.

### 1.3.4 Comparison of Thermoelectric and Thermionic Cooling

Much work has been carried out in comparing thermionic and thermoelectric solid-state coolers. Nolas and Goldsmid [28] compared thermionic and thermoelectric coolers under 'idealised' conditions and showed that thermionic refrigeration is generally superior to thermoelectric refrigeration, but that the materials needed to produce the devices studied in their work are not yet available. Ulrich, Barnes and Vining [29] showed that the performance of both thermoelectric and thermionic coolers relies on the same material parameters. They showed that for all known materials single-barrier thermionic coolers are less effective than thermoelectric coolers.

## 1.4 Thesis Outline

This thesis aims to continue the research carried out in using semiconductor-based thermionic devices to achieve cooling at room temperature. Rather than investigating hybrid systems such as optical-thermionic, thermionic-thermoelectric or thermionic-vacuum devices, this work concentrates on using simple semiconductor multiple-barrier devices. Although hybrid systems may ultimately prove to be more promising, it is believed there is still benefit to be gained from studying straight forward semiconductor multiple-barrier systems.

Chapter 2 gives an introduction to and derivation of Richardson's equation and the accompanying energy transport associated with the electric current. Chapter 3 extends Richardson's equation to develop a simple system that includes two thermionic interfaces. The performance of single-barrier devices is examined and then compared to that of multiple-barrier devices investigated in chapter 4. A numerical solution is developed in this chapter to model $N$-barrier devices. Chapter 5 discusses device design, experimental setup and results. Chapter 6 compares the theoretical $I$-$V$ characteristics generated in

chapter 4 with those obtained experimentally and presented in chapter 5. It is found that the simple model derived does not agree well with the experimental results. An attempt at including non-ideal effects such as space-charge accumulation in the devices into the simple model shows that a more complex model is necessary in modelling the devices. The remainder of chapter 6 deals with developing a more complex model for the devices. Chapter 7 compares the theoretical results of this new model with the experimental results with more success. This new tool is then used to design a new generation of devices and to compare the experimental results from these new devices again with the theoretical model. The poor performance of GaAs-based systems is discussed and devices using different materials are proposed. Coupled with other design parameters that aim to reduce heat transfer due to conduction, these next generation devices are expected to perform much better than the devices discussed in this thesis.

# CHAPTER 2

# THE RICHARDSON-DUSHMAN EQUATION

The following sections discuss the Richardson-Dushman equation and how it applies to solid-state thermionic refrigeration.

## 2.1 A Brief History of Richardson's Equation

The Richardson-Dushman equation (also know simply as Richardson's equation) describes the maximum (or *saturation*) current emitted thermionically from a hot surface. Richardson deduced an equation in 1901 for the *saturation emission current* per unit area of an emitting surface as a function of temperature [30]:

$$I_s = A_1 T^{\frac{1}{2}} \exp\left(-\frac{w}{k_B T}\right) \tag{2.1}$$

where $A_1$ = constant

$w$ = work per electron needed to escape surface (assumed constant for any given material)

$T$ = absolute temperature

$k_B$ = Boltzmann's constant

In formulating this equation Richardson assumed that electrons inside a conductor obey the laws of a perfect gas and share the heat energy with the atoms in the body. In doing so, he incorrectly assumed that electrons follow a *Maxwellian* velocity distribution. This assumption was due to uncertainty at the time about the behaviour of electrons inside conductors.

Richardson later derived an equation using thermodynamic reasoning alone which did not require any assumptions about the velocity distribution of electrons inside conductors [31]:

$$I_s = A T^2 \exp\left(-\frac{\phi_0}{k_B T}\right) \tag{2.2}$$

This equation describes the maximum emission current per unit area from a hot surface. Emitted electrons that are just outside the body and not removed by some means, such as an electric field, will set up a retarding force for the emission of more electrons. This will reduce the emission current to some value less than this saturation value.

Richardson refers to $\phi$ as the change in energy of a system which accompanies the transference of each electron from a hot body to the surroundings.

## 2.2 A Statistical Approach to Richardson's Equation

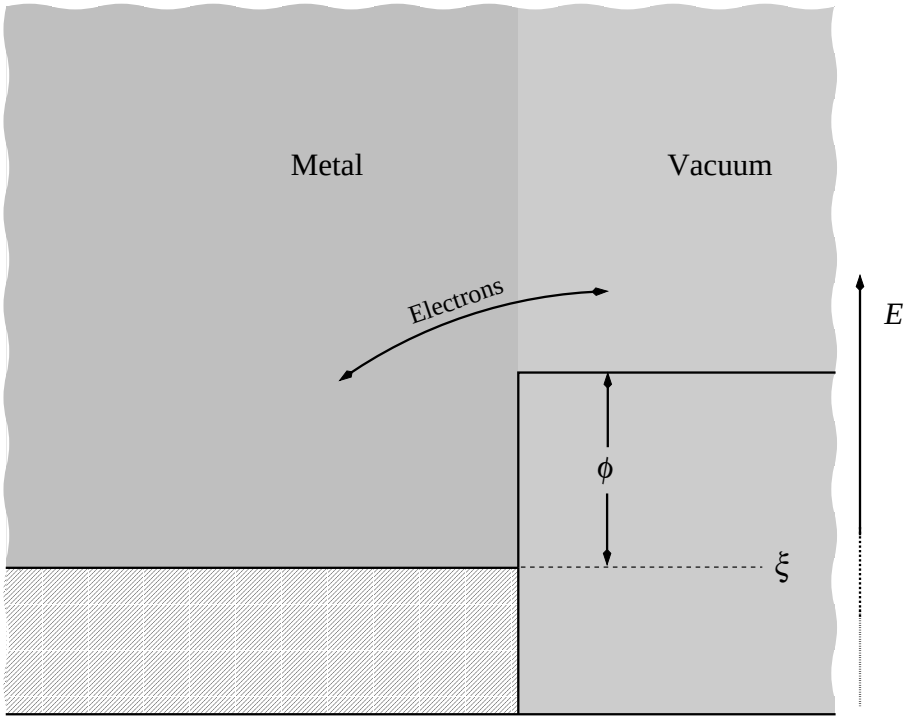


Figure 2.1: Metal-Vacuum Interface

Referring to Fig. 2.1, it is a straight forward procedure to derive Richardson's equation using statistical mechanics. The figure shows a metal-vacuum interface but is just as applicable to the interface between two semiconductor materials. Electrons follow the Fermi-Dirac distribution:

$$f(E) = \frac{1}{\exp\left[(E-\xi)/\left(k_B T\right)\right] + 1} \tag{2.3}$$

Here, $\xi$ denotes the *chemical potential* of the material. This parameter is temperature dependent. The *Fermi energy*, $E_f$, refers to the chemical potential at zero temperature $(E_f = \xi(T = 0))$.

To simplify equations we approximate the Fermi-Dirac distribution by the Maxwell-Boltzmann distribution by ignoring the '1' in the denominator. The validity of doing so is discussed in section 2.3.

The kinetic energy of electrons (the average energy electrons have above the chemical potential) is given by:

$$E - \xi = \frac{m^*}{2}\left(v_x^2 + v_y^2 + v_z^2\right) \tag{2.4}$$

To calculate the current density emitted thermionically from the metal surface, we start with the concentration of electrons whose velocities fall between $(v_x, v_y, v_z)$ and $(v_x + dv_x, v_y + dv_y, v_z + dv_z)$ [32]:

$$d^3N_q = N_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{3}{2}} \exp\left[-\frac{m^*\left(v_x^2 + v_y^2 + v_z^2\right)}{2k_B T}\right] dv_x dv_y dv_z \tag{2.5}$$

where $N_q$ is the quantum concentration of electrons. The term $\left[m^*/\left(2\pi k_B T\right)\right]^{\frac{3}{2}}$ is a normalisation factor to make Eq. 2.5 equal $N_q$ when integrated over all energies.

The $x$-direction is taken perpendicular to the interface, and the current density in the $x$-direction is given by:

$$\begin{aligned} dJ_x &= -qv_x d^3N_q \\ J_x &= \int dJ_x \\ &= -qN_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{3}{2}} \int\int\int v_x \exp\left[-\frac{m^*\left(v_x^2 + v_y^2 + v_z^2\right)}{2k_B T}\right] dv_x dv_y dv_z \end{aligned} \tag{2.6}$$

When integrating over each of the directions, only electrons with enough kinetic energy (and corresponding velocity $v_{x0}$) in the $x$-direction to overcome the potential barrier $q\phi$ contribute to the current density. This is given by:

$$\frac{m^* v_x^2}{2} \geq q\phi$$

so that

$$v_{x0} = \left[\frac{2q\phi}{m^*}\right]^{\frac{1}{2}} \tag{2.7}$$

This sets the limits of integration to $v_{x0} < v_x < \infty$ and $-\infty < v_{y,z} < \infty$, so that:

$$J_x = -qN_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{1}{2}} \int_{v_{x0}}^{\infty} v_x \exp\left(\frac{-m^* v_x^2}{2k_B T}\right) dv_x$$

Then using the fact that:

$$N_q = \left(\frac{m^* k_B T}{2\pi\hbar^2}\right)^{\frac{3}{2}}$$

we arrive at the common form of Richardson's equation:

$$J_x = AT^2 \exp\left(\frac{-q\phi}{k_B T}\right) \tag{2.8}$$

where

$$A = \frac{m^* q k_B^2}{2\pi^2\hbar^3} \approx 1.2 \times 10^6 \frac{\text{A}}{\text{m}^2\text{K}^2} \tag{2.9}$$

## 2.3 Validity of Using Boltzmann Distribution for Low Work Function or High Temperature Semiconductor Barrier

When considering thermionic refrigeration it is usual practice to consider that Richardson's equation is valid for the electrons given off from a semiconductor barrier. The assumption made is that the region of emission is in the tail of the energy distribution of carriers, shown in Fig. 2.2.

When there is a large work function, $q\phi$, or small temperature, the tail of the two distributions align sufficiently that the Maxwell-Boltzmann distribution becomes a good approximation for the Fermi-Dirac distribution and can therefore be used in its place. From an analytical point of view this is desirable as integration of the Fermi-Dirac function becomes cumbersome. The whole concept of semiconductor thermionic refrigeration being efficient relies on the use of materials with small work functions [10]. It is therefore necessary to determine whether Richardson's equation for thermionic emission of electrons is valid in devices with relatively small functions operating near room temperature. To do so, numerical methods are employed to compare the distribution functions.

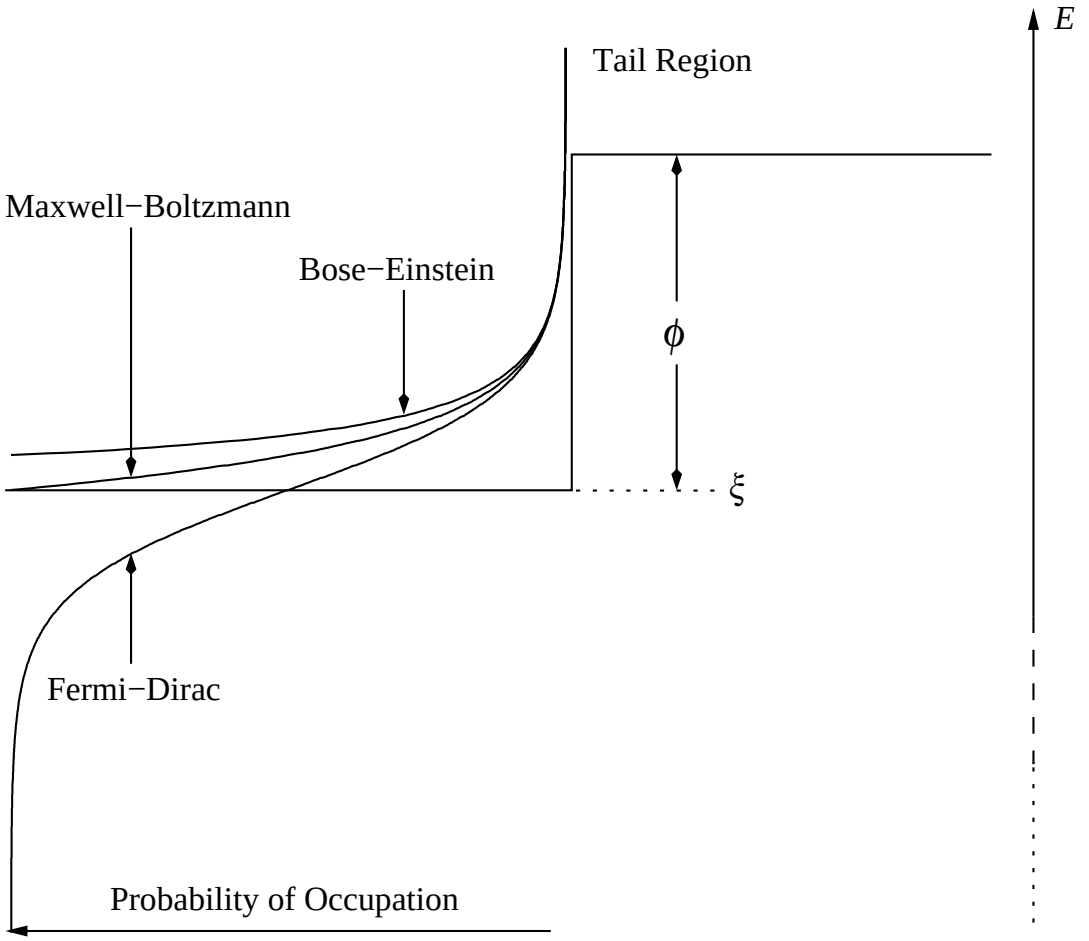


Figure 2.2: Comparison of Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac Distributions

### 2.3.1 The Equations

The Maxwell-Boltzmann distribution is given by:

$$f(E) = \frac{1}{\exp\left[(E-\xi)\,/\,(k_B T)\right]}, \tag{2.10}$$

while the Fermi-Dirac distribution is given by:

$$f(E) = \frac{1}{\exp\left[(E-\xi)\,/\,(k_B T)\right] + 1} \tag{2.11}$$

and the Bose-Einstein distribution is given by:

$$f(E) = \frac{1}{\exp\left[(E-\xi)\,/\,(k_B T)\right] - 1} \tag{2.12}$$

By referring to Fig. 2.2 it can be seen that all three distributions approach each other in the tail region. As long as the emission region, $(\xi + q\phi)$, is sufficiently in the tail region then the Maxwell-Boltzmann distribution can successfully be used in place of either the

Fermi-Dirac or Bose-Einstein distribution.

Put another way, it is sufficient to use the Maxwell-Boltzmann distribution in calculations when the devices are operating in 'classical' mode rather than a regime where quantum effects become important. This occurs when

$$\exp\left(\frac{E-\xi}{k_BT}\right) \gg 1 \tag{2.13}$$

so that the $\pm 1$ term in the Fermi-Dirac and Bose-Einstein distributions becomes negligible. In the case of thermionic current, the only electrons involved in emission are ones with enough kinetic energy to overcome the potential barrier imposed by $q\phi$. The electrons with least energy involved in producing current are the ones with just enough energy to overcome the potential barrier, $E_{min} = \xi + q\phi$. For the Maxwell-Boltzmann function to be a valid substitution for the Fermi-Dirac function with the relative error less than 5% for all electrons involved in current flow it can be shown that $E - \xi = q\phi > 3k_BT$, using:

$$\left(\frac{1}{A} - \frac{1}{A+1}\right) / \left(\frac{1}{A+1}\right) < 0.05$$

where $A = \exp\left(\frac{E-\xi}{k_BT}\right) = \exp\left(\frac{q\phi}{k_BT}\right)$

Solving for A:

$$\frac{1}{A} \;<\; 0.05 \Rightarrow A > 20$$

and then $q\phi$:

$$\begin{aligned} \exp\left(\frac{q\phi}{k_BT}\right) &> 20 \\ \phi &> \frac{k_BT}{q}\ln 20 \\ \text{So}\quad \phi &> \frac{3k_BT}{q} \end{aligned}$$

At room temperature ($\sim$300 K) this means that $\phi > 77.6$ mV. When a full integration with respect to energy is carried out using Maxwell-Boltzmann statistics this error reduces further because as the energy of the electrons increases the Maxwell-Boltzmann distribution approaches the Fermi-Dirac distribution even more closely.

This analysis shows that by only considering devices in which the barrier height is larger than around 77 mV the Maxwell-Boltzmann distribution can be used with minimal error.

The next section derives the energy transport associated with the electric current due to thermionic emission.

## 2.4 Derivation of Thermionic Heat Current using Maxwell-Boltzmann statistics

Richardson's equation describes the current density emitted from a surface or interface due to thermionic emission. To investigate the possibility of using a semiconductor-based system to achieve cooling it is necessary to determine the amount of energy removed by this current flow. This section derives an equation to describe this energy transfer. In order to simplify calculations Maxwell-Boltzmann statistics are used as an approximation for Fermi-Dirac statistics.

Starting off with the expression for thermionic electrical current:

$$J = -qN_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{3}{2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{v_{xo}}^{\infty} \frac{v_x dv_x dv_y dv_z}{\exp\left[m^*\left(v_x^2+v_y^2+v_z^2\right)/(2k_B T)\right]} \tag{2.14}$$

we take the average energy of each transmitted electron, $q\phi_t$, to be a function of velocity (kinetic energy) in the direction of current flow, in this case the $x$-direction:

$$q\phi_t\left(v_x\right) = \frac{m^* v_x^2}{2} \tag{2.15}$$

where $\phi_t$ is in volts.

We then have:

$$J_q = -qN_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{3}{2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{v_{xo}}^{\infty}\left(\frac{m^* v_x^2}{2q}+\frac{k_B T}{q}\right)\frac{v_x dv_x dv_y dv_z}{\exp\left[m^*\left(v_x^2+v_y^2+v_z^2\right)/(2k_B T)\right]}$$

$$= -N_q\left(\frac{m^*}{2\pi k_B T}\right)^{\frac{3}{2}}\left[k_B T\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underbrace{\int_{v_{xo}}^{\infty}\frac{v_x dv_x dv_y dv_z}{\exp\left[m^*\left(v_x^2+v_y^2+v_z^2\right)/(2k_B T)\right]}}_{A}\right.$$

$$\left. + \frac{m^*}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underbrace{\int_{v_{xo}}^{\infty}\frac{v_x^3 dv_x dv_y dv_z}{\exp\left[m^*\left(v_x^2+v_y^2+v_z^2\right)/(2k_B T)\right]}}_{B}\right]$$

where $v_{xo}^2 = 2q\phi/m^*$

Solving A:

$$A = \exp\left[-m^*\left(v_y^2+v_z^2\right)/(2k_B T)\right]\int_{v_{xo}}^{\infty} v_x \exp\left(-\frac{m^* v_x^2}{2k_B T}\right)dv_x$$

$$\text{let } u = \frac{m^* v_x^2}{2k_B T}\text{ , } du = \frac{m^* v_x}{k_B T}dv_x$$

$$\text{so } v_x dv_x = \frac{k_B T}{m^*}du$$

Then:

$$A = \frac{k_B T}{m^*}\exp\left[-\frac{m^*}{2k_B T}\left(v_y^2+v_z^2\right)\right]\int_{\frac{m^* v_{xo}^2}{2k_B T}}^{\infty}\exp(-u)du$$

$$= \frac{k_B T}{m^*}\exp\left[-\frac{m^*}{2k_B T}\left(v_y^2+v_z^2\right)\right]\exp\left[-\frac{m^* v_{xo}^2}{2k_B T}\right]$$

Solving B:

$$B = \exp\left[-\frac{m^*}{2k_B T}\left(v_y^2+v_z^2\right)\right]\int_{v_{xo}}^{\infty} v_x^3 \exp\left(-\frac{m^* v_x^2}{2k_B T}\right)dv_x$$

$$\text{let } u = \frac{m^* v_x^2}{2k_B T}\text{ , } du = \frac{m^* v_x}{k_B T}dv_x$$

$$\text{so } v_x^3 dv_x = 2\left(\frac{k_B T}{m^*}\right)^2 u du$$

Then:

$$B = 2\left(\frac{k_B T}{m^*}\right)^2\exp\left[-\frac{m^*}{2k_B T}\left(v_y^2+v_z^2\right)\right]\int_{\frac{m^* v_{xo}^2}{2k_B T}}^{\infty} u\exp(-u)du$$

Using integration by parts:

$$B = 2\left(\frac{k_BT}{m^*}\right)^2\left(\frac{m^*v_{xo}^2}{2k_BT}+1\right)\exp\left[-\frac{m^*}{2k_BT}\left(v_y^2+v_z^2\right)\right]\exp\left[-\frac{m^*v_{xo}^2}{2k_BT}\right]$$

Combining the results for $A$ and $B$ we get:

$$\begin{aligned} J_q &= -N_q\left(\frac{m^*}{2\pi k_BT}\right)^{\frac{3}{2}}\left[k_BT+k_BT\left(\frac{m^*v_{xo}^2}{2k_BT}+1\right)\right] \\ &\times \exp\left[-\frac{m^*v_{xo}^2}{2k_BT}\right]\frac{k_BT}{m^*}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dv_ydv_z}{\exp\left[m^*\left(v_y^2+v_z^2\right)/\left(2k_BT\right)\right]} \end{aligned}$$

It can be shown that:

$$J = -qN_q\left(\frac{m^*}{2\pi k_BT}\right)^{\frac{3}{2}}\exp\left[-\frac{m^*v_{xo}^2}{2k_BT}\right]\frac{k_BT}{m^*}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dv_ydv_z}{\exp\left[m^*\left(v_y^2+v_z^2\right)/\left(2k_BT\right)\right]}$$

so that:

$$J_q = \left[\frac{k_BT}{q}+\frac{k_BT}{q}\left(\frac{m^*v_{xo}^2}{2k_BT}+1\right)\right]J$$

but:

$$\frac{m^*v_{xo}^2}{2k_BT} = \frac{q\phi}{k_BT}$$

giving:

$$\begin{aligned} J_q &= \left[\frac{k_BT}{q}+\frac{k_BT}{q}\left(\frac{q\phi}{k_BT}+1\right)\right]J \\ &= \left(\phi+\frac{2k_BT}{q}\right)J \end{aligned}$$

In this chapter expressions have been derived for the electrical and energy current densities emitted thermionically from surfaces. Because the devices investigated in this work have uniform cross-sectional areas, the term current will be used interchangeably for current density.

The next chapter will build on the results of this chapter to develop equations describing the electrical and accompanying energy current in a single-barrier semiconductor thermionic device.

# CHAPTER 3

# THERMIONIC REFRIGERATION IN A SINGLE-BARRIER SEMICONDUCTOR DEVICE

## 3.1 Building a Simple Model

The simplest model for thermionic emission in a single-barrier semiconductor device can be made by considering two thermionic interfaces back to back. If a single interface is used the heat conduction from the hot-to-cold side will cancel out the energy taken from the cold-to-hot side. No net cooling will be achieved. By using two interfaces with a finite barrier in between, the amount of conduction from the hot to cold side can be reduced by considering the thermal current proportional to the thermal gradient [33]:

$$J_q^{cond} = \frac{\Delta T}{R^{th}} \tag{3.1}$$

where $R^{th}$ is the thermal resistance of the barrier. If the heat transport from the cold side due to thermionic emission is greater than that due to conduction from the hot side then, neglecting any internal heat generation, net cooling will be achieved.

In order for the shape of the barrier to be inconsequential (and for this simple approach to be valid), no electrons involved in transport should interact with the material in the barrier. To achieve this the electrons must traverse the barrier material *ballistically*. For this to happen the barrier width must be of the order of the *mean free path* of electrons in the material. Because the thermal resistance of the barrier is dependent on it's width, this sets a maximum for the thermal resistance between electrodes which, in turn, sets an upper limit for the maximum thermal efficiency possible for a given device. A full discussion of thermal resistance and barrier width is given in section 3.5.1 but for the moment we can consider the thermal resistance to simply be a controllable parameter of

the device.

### 3.1.1 Assumptions relating to an 'ideal' thermionic device

The development of this model allows a great deal to be learnt about the device, even if it is not completely accurate. There are a number of important assumptions that have to be made when using this simple approach. Indeed the usefulness of using this approach will be shown in later sections to be quite limited. Nevertheless, this method provides a starting point from which it will be possible to actually determine how well the model describes a real system.

The assumptions that this model makes include:

- There is no space charge accumulation in the device;
- Any potential applied across the device is dropped only in the barrier itself and not in any of the electrodes – this can be approximated by doping the electrodes much more than the barrier;
- Carriers contributing to current flow travel ballistically through the barrier and so do not interact at all in the barrier – the barrier width is less than the mean free path of carriers in the barrier material.
- Following from the previous point, the shape of the barrier is inconsequential;
- The barrier width is greater than the tunnelling width of carriers in the barrier;
- Richardson's equation is for a system in equilibrium, so our system is assumed to be in, or very close to, equilibrium.

## 3.2 Electrical and heat currents in the device

Fig. 3.1 shows the schematic of a single-barrier thermionic device. For this device, a positive bias of $V$ (V) has been applied to electrode 2 with respect to electrode 1. Electrode 1 is at temperature $T_c$ (K) $< T_h$ (K) , the temperature of electrode 2. We assume

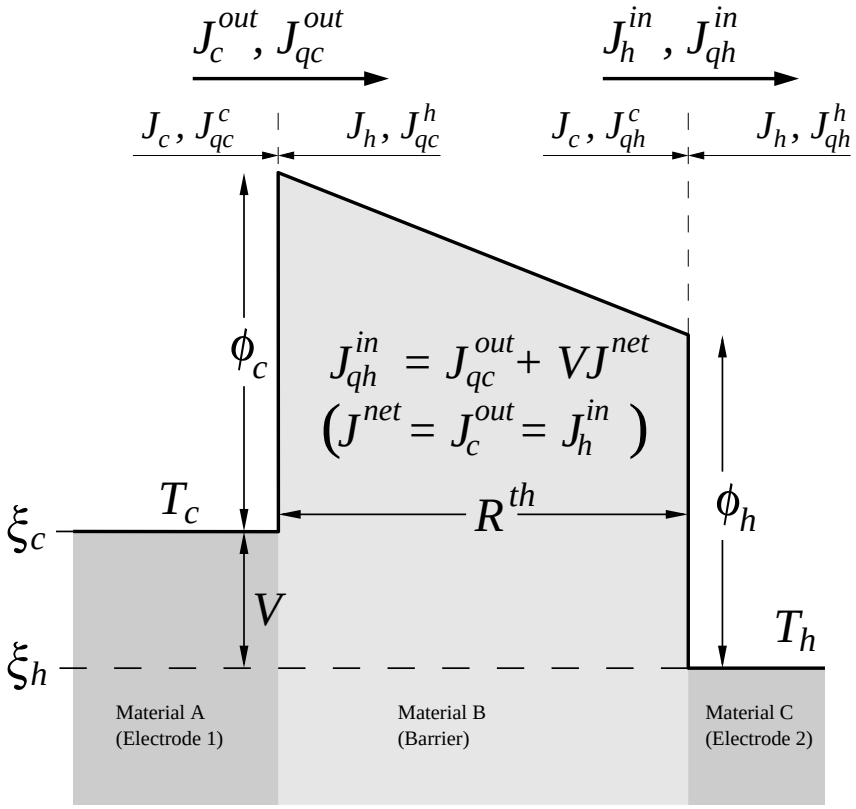


Figure 3.1: Schematic of Single-Barrier Thermionic Device Under Applied Bias $V$

that a net electrical and heat current leaves electrode 1 and moves from left to right. A thermal resistance, $R^{th}$, exists between the two electrodes. A potential barrier, $\phi_c$, exists between electrode 1 and the barrier. A potential barrier, $\phi_h$, exists between the barrier and electrode 2. If material 'A' is identical to material 'C' then $\phi_c = \phi_h$.

### 3.2.1 Net electrical current leaving electrode 1

Using this model it is easy to calculate the net electrical currents leaving electrode 1 and travelling across the barrier:

$$J^{net} = J_c - J_h \tag{3.2}$$

An expression for $J_c$ can be taken from Eq. 2.8. Although this equation was derived for a metal-vacuum interface, the same form can be used in a semiconductor-semiconductor interface by simply changing the effective mass of electrons in Richardson's constant. When the effective mass is different between the two materials, the smaller effective

mass is to be used, as shown by Shakouri and Bowers [12] and Grinberg [34].

Following the derivation in section 2.2, the thermionic electric current that leaves Electrode 1 (where a potential barrier $\phi_c$ is present) is given by:

$$J_c = AT_c^2 \exp\left(-\frac{q\phi_c}{k_B T_c}\right) \tag{3.3}$$

In the configuration given in Fig. 3.1, an electron sitting at the chemical potential of electrode 2 needs to overcome a potential energy of $q\left(\phi_c + V\right)$ in order to contribute to current flow. This then becomes the barrier height, or work function, for electrons leaving electrode 2:

$$J_h = AT_h^2 \exp\left(-\frac{q\left(\phi_c + V\right)}{k_B T_h}\right) \tag{3.4}$$

It is interesting to note that the size of the barrier $\phi_h$ does not influence the magnitude of electrical current that flows in the device. This statement is valid as long as $\phi_c + V \geq \phi_h$. If this condition is not satisfied, then the operating mode of the device changes and care must be take when modelling such devices. This will be dealt with further in section 3.4 for single-barrier devices and chapter 4 for multi-barrier devices.

Combining Eq. 3.3 and Eq. 3.4 we get:

$$\begin{aligned} J^{net} &= J_c - J_h \\ &= AT_c^2 \exp\left(-\frac{q\phi_c}{k_B T_c}\right) - AT_h^2 \exp\left(-\frac{q\left(\phi_c + V\right)}{k_B T_h}\right) \end{aligned} \tag{3.5}$$

If we ignore internal heat generation and carrier collisions in the device then, because electrical current is independent of individual electron energies, this electrical current is constant throughout the device. In other words, the net electrical current leaving electrode 1 to the right is exactly equal to the net current entering electrode 2 from the left.

The condition for which a positive current leaves the cold electrode to the right is given by:

$$V > V_0^J = \frac{2k_B T_h}{q} \log\left(\frac{T_h}{T_c}\right) + \phi_c \left(\frac{T_h}{T_c} - 1\right) \tag{3.6}$$

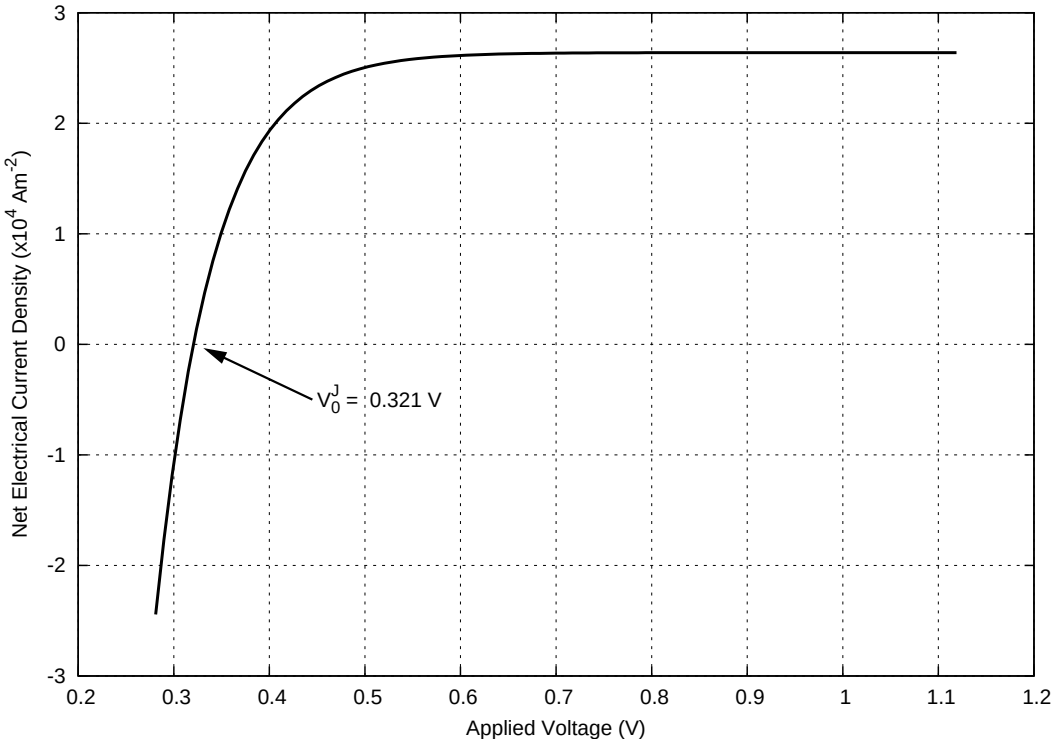


Figure 3.2: Current Density of Single-Barrier Device showing switch in sign at applied bias $V_0^J$ given by Eq. 3.6 ($\phi_c = 0.7$ eV, $T_c = 500$ K, $T_h = 700$ K).

If the applied voltage is less than this then the larger thermal energy of electrons on the right hand side of the barrier causes a net current to enter the cold electrode. The current density versus applied bias for a single-barrier device with $\phi_c = 0.7$ eV, $T_c = 500$ K and $T_h = 700$ K is shown in Fig. 3.2. The applied bias calculated using Eq. 3.6 is shown on the plot.

It can be seen that the net electrical current *saturates* at an applied bias of around 0.7 V. This occurs because the exponential term in $J_h$ becomes so large that no current flows into the cold electrode from the hot electrode. The saturation current is therefore simply:

$$J^{sat} = J_c \tag{3.7}$$

which has a value of $2.64 \times 10^4$ Am$^{-2}$ for the device being examined.

### 3.2.2 Net heat current leaving electrode 1 ignoring conduction heat current

The energy current, as opposed to the electrical current, is not constant throughout the device. The energy current leaving electrode 1 to the right is given by:

$$J_{qc}^{c} = \left(\phi_c + \frac{2k_B T_c}{q}\right) J_c \tag{3.8}$$

The energy current entering electrode 1 from the right is given by:

$$J_{qc}^{h} = \left(\phi_c + \frac{2k_B T_h}{q}\right) J_h \tag{3.9}$$

Each electron travelling from the right deposits, on average, a potential energy of $q\phi_c$ in addition to a kinetic component given by the average thermal energy, $2k_B T_h$.

Combining Eq. 3.8 and Eq. 3.9 the net heat current leaving electrode 1 is given by:

$$\begin{aligned} J_{qc}^{out} &= J_{qc}^{c} - J_{qc}^{h} \\ &= \left(\phi_c + \frac{2k_B T_c}{q}\right) J_c - \left(\phi_c + \frac{2k_B T_h}{q}\right) J_h \end{aligned} \tag{3.10}$$

As for electrical currents, there is a critical voltage at which the net heat current leaving the cold electrode becomes positive. This is the region that we are interested in as a refrigerator. Due to the larger prefactor of $J_h$ in Eq. 3.10, namely $(\phi_c + 2k_B T_h/q)$, than that of $J_c$, $(\phi_c + 2k_B T_c/q)$, this voltage is always larger than the critical bias, $V_0^J$, given in Eq. 3.6:

$$\begin{aligned} V > V_0^{Jq} &= \frac{2k_B T_h}{q} \log\left(\frac{T_h}{T_c}\right) + \phi_c \left(\frac{T_h}{T_c} - 1\right) + \frac{k_b T_h}{q} \log\left(\frac{q\phi_c + 2k_B T_h}{q\phi_c + 2k_B T_c}\right) \\ &= V_0^J + \frac{k_b T_h}{q} \log\left(\frac{q\phi_c + 2k_B T_h}{q\phi_c + 2k_B T_c}\right) \end{aligned} \tag{3.11}$$

The energy current density for the device using the previously defined parameters is shown in Fig. 3.3. The bias at which the heat current becomes positive is 0.323 V, given by Eq. 3.11. Similar to the electrical current, the heat current also saturates:

$$J_{qc}^{sat} = J_{qc}^{c} \tag{3.12}$$

For this device the saturation energy current density is $2.07 \times 10^4$ Wm$^{-2}$.

For the device parameters used in the above example, the two critical voltages, $V_0^{Jq}$ and $V_0^J$, are quite close. For other device parameters the distance between the two critical

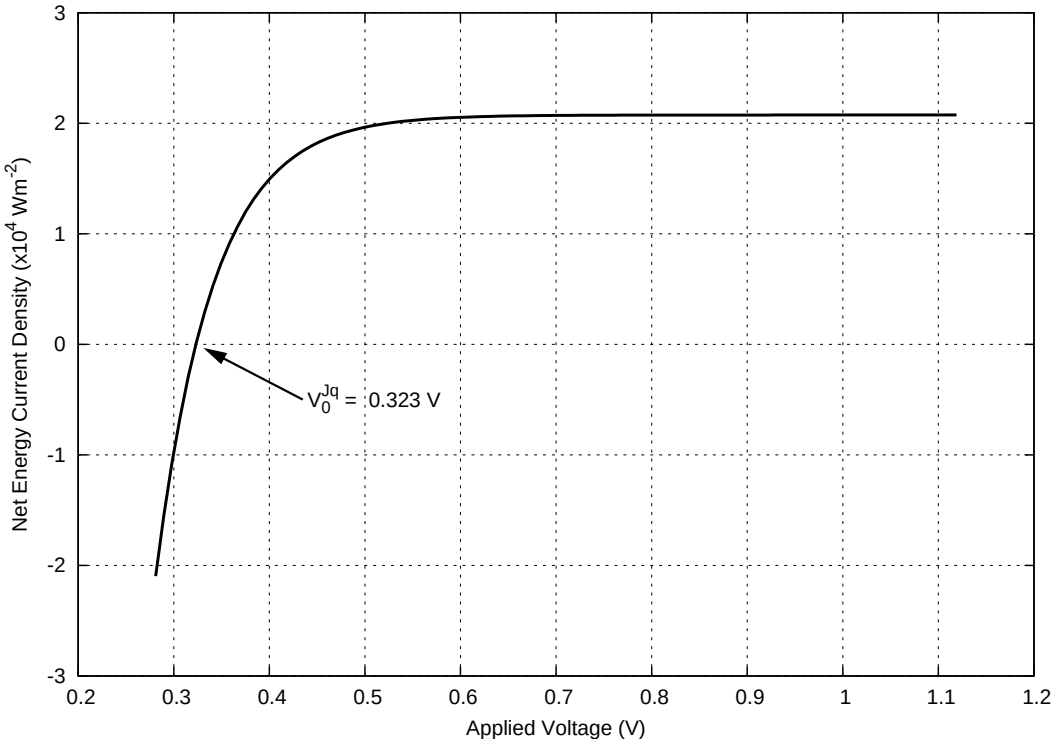


Figure 3.3: Energy Current Density of Single-Barrier Device showing switch in sign at applied bias $V_0^{Jq}$ given by Eq. 3.11 ($\phi_c = 0.7$ eV, $T_c = 500$ K, $T_h = 700$ K).

voltages can increase quite substantially. For a given hot and cold temperature, this difference will increase as the barrier height of the cold electrode decreases. This means that a greater amount of work must be input into the system before heat starts to be removed from the cold side.

### 3.2.3 Net heat current leaving electrode 1 including conduction heat current

Incorporating heat current due to conduction, $\Delta T/R^{th}$, into the model gives us a model built using simple assumptions:

$$J_{qc}^{out} = \left(\phi_c + \frac{2k_BT_c}{q}\right) J_c - \left(\phi_c + \frac{2k_BT_h}{q}\right) J_h - \frac{\Delta T}{R^{th}} \tag{3.13}$$

This conduction heat current always works to reduce the net heat current leaving electrode 1 thus reducing the overall performance of the device. The main objective of finding a useful device is one in which this heat current is reduced as much as possible while still satisfying conditions for ballistic transport of carriers.

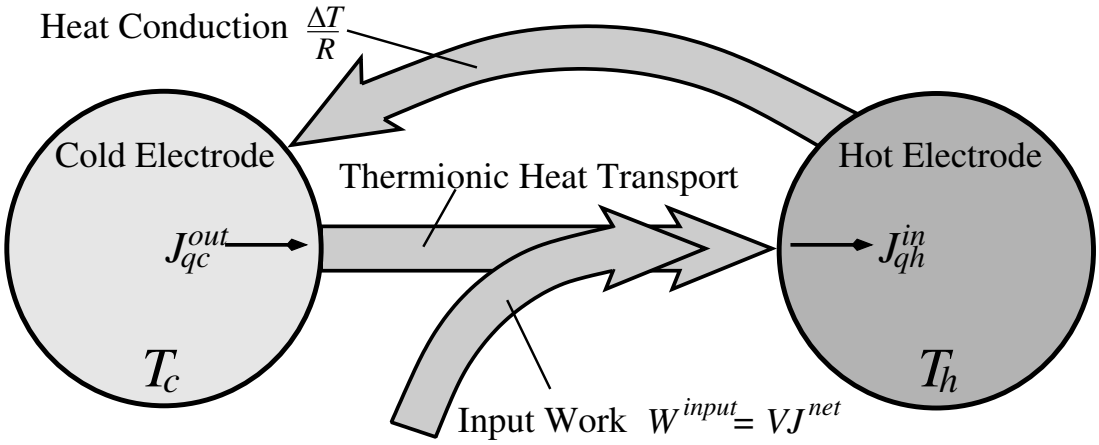


Figure 3.4: Thermodynamic representation of thermionic device

### 3.2.4 Net heat current entering electrode 2 including conduction heat current

With respect to electrode 2, the net heat current entering from the left is given by:

$$J_{qh}^{in} = \left(\phi_c + V + \frac{2k_B T_c}{q}\right) J_c - \left(\phi_c + V + \frac{2k_B T_h}{q}\right) J_h - \frac{\Delta T}{R^{th}} \tag{3.14}$$

This expression is exactly the same as that given by Eq. 3.13 except that the potential energy contribution of each electron is increased by $V$ because of the applied bias. Potential energy in the hot barrier is measured from its chemical potential, which has been lowered by an amount $V$ with respect to the chemical potential of the cold side.

### 3.2.5 Work input into the system

Fig. 3.4 shows a thermodynamic representation of a thermionic device. This figure includes heat currents and work input into the system. The heat current leaving the cold reservoir is $J_{qc}^{out}$ and the heat current entering the hot reservoir is $J_{qh}^{in} = J_{qc}^{out} + W^{input}$. The work input into the system is then the difference between these two heat currents:

$$\begin{aligned} W^{input} &= J_{qh}^{in} - J_{qc}^{out} \\ &= \left[\left(\phi_c + V + \frac{2k_B T_c}{q}\right) J_c - \left(\phi_c + V + \frac{2k_B T_h}{q}\right) J_h - \frac{\Delta T}{R^{th}}\right] \\ &- \left[\left(\phi_c + \frac{2k_B T_c}{q}\right) J_c - \left(\phi_c + \frac{2k_B T_h}{q}\right) J_h - \frac{\Delta T}{R^{th}}\right] \\ &= V(J_c - J_h) \end{aligned}$$

$$= \quad VJ^{net} \tag{3.15}$$

which is the electrical power put into the system. Also shown in the figure is the parasitic component of energy current due to conduction which reduces overall device performance.

## 3.3 Device Performance

### 3.3.1 Thermal Efficiency

Equations 3.5, 3.13 and 3.15 allow us to calculate the thermal efficiency or *coefficient of performance* (COP) of the single-barrier device:

$$\eta = \frac{J_{qc}^{out}}{W^{input}} = \frac{J_{qc}^{out}}{VJ^{net}} \tag{3.16}$$

with both electrical and energy currents considered to be positive when moving to the right. The COP is simply a measure of the amount of heat energy removed from the cold surface compared to the amount of input energy needed to do so. Because these equations are so simple, they can be solved directly. Using device parameters $\phi_c = 0.7$ eV, $T_c = 500$ K and $T_h = 700$ K the efficiency of a single barrier device (neglecting conduction heat current for the time being) is shown in Fig. 3.5. The same figure is shown in the paper by Mahan [10].

The maximum theoretical efficiency for any reversible refrigerator is given by [35]:

$$\eta_{max} = \frac{T_c}{T_h - T_c} = \frac{T_c}{\Delta T} \tag{3.17}$$

so for the device operating between 700 K and 500 K the maximum theoretical efficiency is 2.5. The device in Fig. 3.5 has a maximum efficiency of 2.08 at around 0.35 V. This represents an efficiency that is 83% of the theoretical limit. By modifying device parameters (such as barrier heights) the device efficiency can be increased. This will be dealt with in more detail in later sections.

It is very important to note at this point that the efficiency of the device shown in Fig. 3.5 is calculated with the heat current due to heat conduction neglected (the thermal resistance of the device is infinite). The inclusion of this term reduces the efficiency of devices

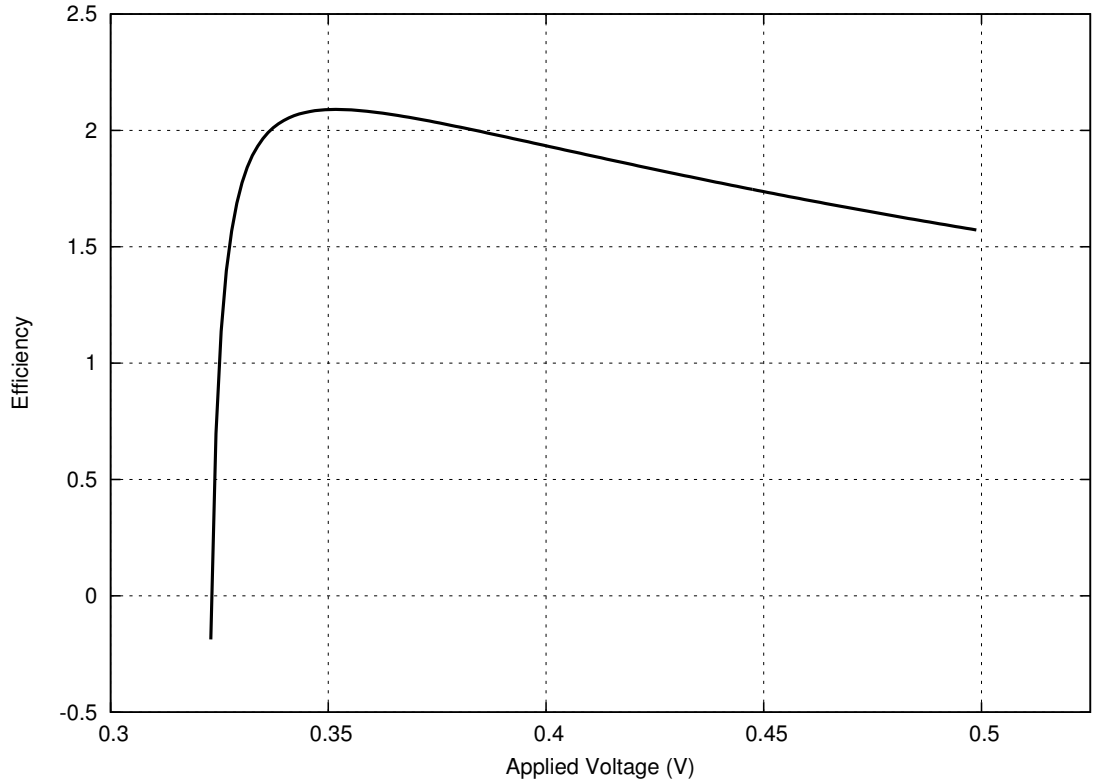


Figure 3.5: Theoretical Efficiency of Single-Barrier Device vs. Applied Voltage ($\phi_c = 0.7$ eV, $T_c = 500$ K, $T_h = 700$ K)

dramatically, as will be seen in later sections, so should not be neglected.

### 3.3.2 Heat current removed from cold side

In addition to trying to maximise device efficiency, another property that must be maximised is the heat current, $J_{qc}^{out}$, that is removed from the cold side of the device. There is little point in having a highly efficient device if it does not remove enough heat from the cold side to achieve appreciable cooling. It will be seen that increasing the heat current removed for a given device reduces the efficiency, but this is true for any reversible refrigerator.

## 3.4 Device Classes and Operating Modes

The following sections discuss two different classes of single-barrier device. For both cases it is assumed that the applied bias is positive. A negative voltage would increase heat current from the hot to cold side meaning that the device is not operating as a refrigerator. As Eq. 3.11 has shown a positive bias can be applied and a net heat current can still flow from the hot to cold side even in the absence of conduction heat current.

Conduction heat current is ignored for the moment (the thermal resistance of devices is considered infinite) as it is a constant term dependent only on $T_c$, $T_h$ and $R^{th}$ which will all be held constant and equal for both classes.

### 3.4.1 Devices having $\phi_c \geq \phi_h$ (Class 1 Devices)

The device shown in Fig. 3.1 is operating with $\phi_c + V \geq \phi_h$. This condition will be referred to as *Mode 1*. This mode of operation can always be ensured by having a class of device in which $\phi_c \geq \phi_h$, operating with $V \geq 0$. This type of device will be referred to as a *Class 1* device. All equations developed so far are *always* valid for this class of device. In fact, Class 1 devices can only operate in Mode 1.

Another class of device is one in which $\phi_h > \phi_c$. The equations developed so far (eg. Eq. 3.5 and Eq. 3.13) are not always valid for this type of device and will be discussed further in the next section.

### 3.4.2 Devices having $\phi_h > \phi_c$ (Class 2 Devices)

Fig. 3.6 shows a device in which $\phi_h > \phi_c$ and $V > 0$. Case a) depicts when $\phi_h > \phi_c + V$. Under this regime equations 3.5 and 3.13 are no longer valid. This mode of operation will be referred to as *Mode 2* and is defined explicitly by a device having $\phi_h > \phi_c + V$ when $V > 0$.

In Mode 2 the net current 'leaving' the cold electrode to the right and moving across the barrier will have a different form to that of Eq. 3.5. The term 'leaving' is used for

consistency when in actual fact the net current can be negative and may enter the cold electrode from the right, depending on the magnitude of the applied bias.

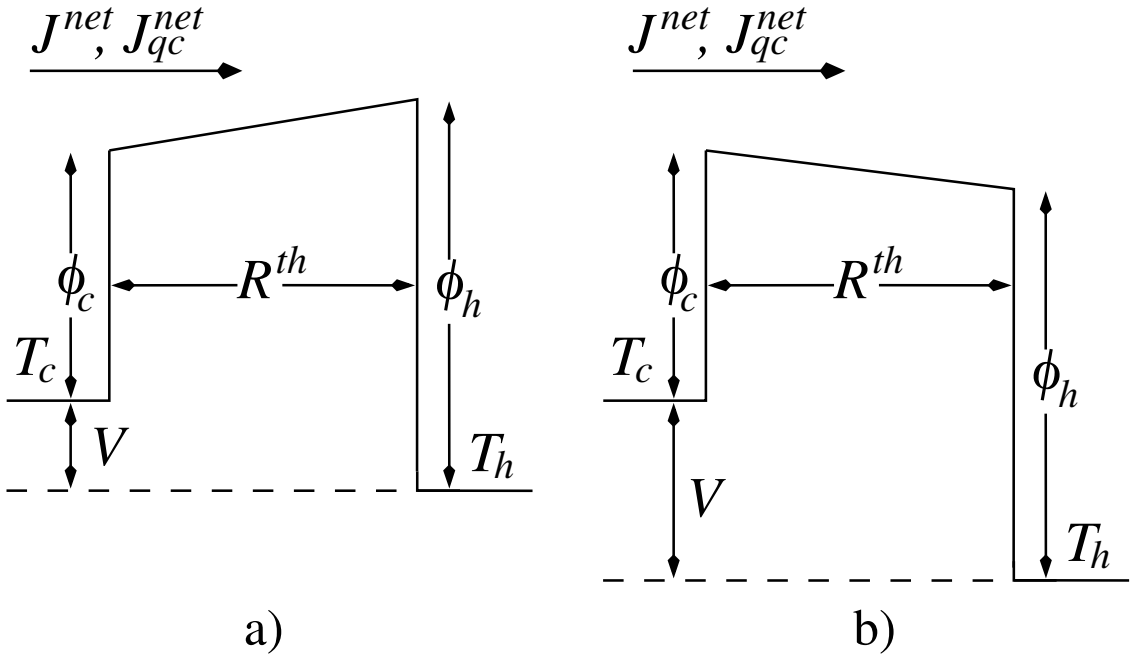


Figure 3.6: Schematic of a Class 2 Single-Barrier Thermionic Device ($\phi_h > \phi_c$) where a) $\phi_h > \phi_c + V$ (Mode 2) and b) $\phi_c + V > \phi_h$ (Mode 1)

The expression for this current is given by:

$$J^{net} = AT_c^2 \exp\left(-\frac{q\left(\phi_h - V\right)}{k_B T_c}\right) - AT_h^2 \exp\left(-\frac{q\phi_h}{k_B T_h}\right) \tag{3.18}$$

Electrons leaving the cold side must overcome a potential of $(\phi_h - V)$ to contribute to current flow, while electrons leaving the hot side must overcome a potential of $\phi_h$. Analogously to the electrical current derived in section 3.2.1, $\phi_c$ does not enter into this equation.

Similarly, the net heat current leaving the cold electrode, ignoring conduction heat current, is given by:

$$\begin{aligned} J^{out}_{qc} &= \left(\phi_h - V + \frac{2k_B T_c}{q}\right) AT_c^2 \exp\left(-\frac{q\left(\phi_h - V\right)}{k_B T_c}\right) \\ &- \left(\phi_h - V + \frac{2k_B T_h}{q}\right) AT_h^2 \exp\left(-\frac{q\phi_h}{k_B T_h}\right) \end{aligned} \tag{3.19}$$

Each electron travelling to the right takes, on average, a potential energy of $q\left(\phi_h - V\right)$

in addition to a kinetic component given by the average thermal energy, $2k_BT_c$. Each electron travelling from the right deposits a potential energy of $q\left(\phi_h - V\right)$ in addition to a kinetic component given by the average thermal energy, $2k_BT_h$.

The condition for when $J^{net} > 0$, i.e. a net electrical current flows to the right, occurs when the applied voltage is larger than:

$$V_{0\ (Class\ 2)}^{J} = \frac{2k_BT_c}{q}\log\left(\frac{T_h}{T_c}\right) + \phi_h\left(1 - \frac{T_c}{T_h}\right) \tag{3.20}$$

Similarly, for positive heat current (heat is being removed from the cold electrode), the applied bias must be larger than:

$$\begin{aligned} V_{0\ (Class\ 2)}^{Jq} &= \frac{2k_BT_c}{q}\log\left(\frac{T_h}{T_c}\right) + \phi_h\left(1 - \frac{T_c}{T_h}\right) \\ &+ \frac{k_BT_c}{q}\log\left(\frac{q\phi_h - qV_{0\ (Class\ 2)}^{Jq} + 2k_BT_h}{q\phi_h - qV_{0\ (Class\ 2)}^{Jq} + 2k_BT_c}\right) \\ &= V_{0\ (Class\ 2)}^{J} + \frac{k_BT_c}{q}\log\left(\frac{q\phi_h - qV_{0\ (Class\ 2)}^{Jq} + 2k_BT_h}{q\phi_h - qV_{0\ (Class\ 2)}^{Jq} + 2k_BT_c}\right) \end{aligned} \tag{3.21}$$

This expression is similar to the one given by Mahan [10] but differs due to sign conventions. As Mahan notes, because $V_{0\ (Class\ 2)}^{Jq}$ appears in the $\log()$ term, this must be solved numerically. This adds to the complexity of solving the equations.

For the device in Fig. 3.6 to have a positive net electrical current and still be operating in Mode 2 ($\phi_h > \phi_c + V$) the following relationship between $\phi_h$ and $\phi_c$ must be satisfied. This is found by substituting $\phi_h > \phi_c + V$ into Eq. 3.20:

$$\phi_h > \frac{2k_BT_c}{q}\log\left(\frac{T_h}{T_c}\right) + \phi_c\frac{T_h}{T_c} \tag{3.22}$$

If this condition is not satisfied then there is no applied bias that will induce a positive net electrical current while still operating in Mode 2. Instead the device will be operating in Mode 1. This mode is shown as b) in Fig. 3.6 and the net electrical and heat currents leaving the cold electrode to the right are given by the previously derived Eq. 3.5 and Eq. 3.10 respectively. Thus, Class 2 devices can operate in both Mode 1 and Mode 2.

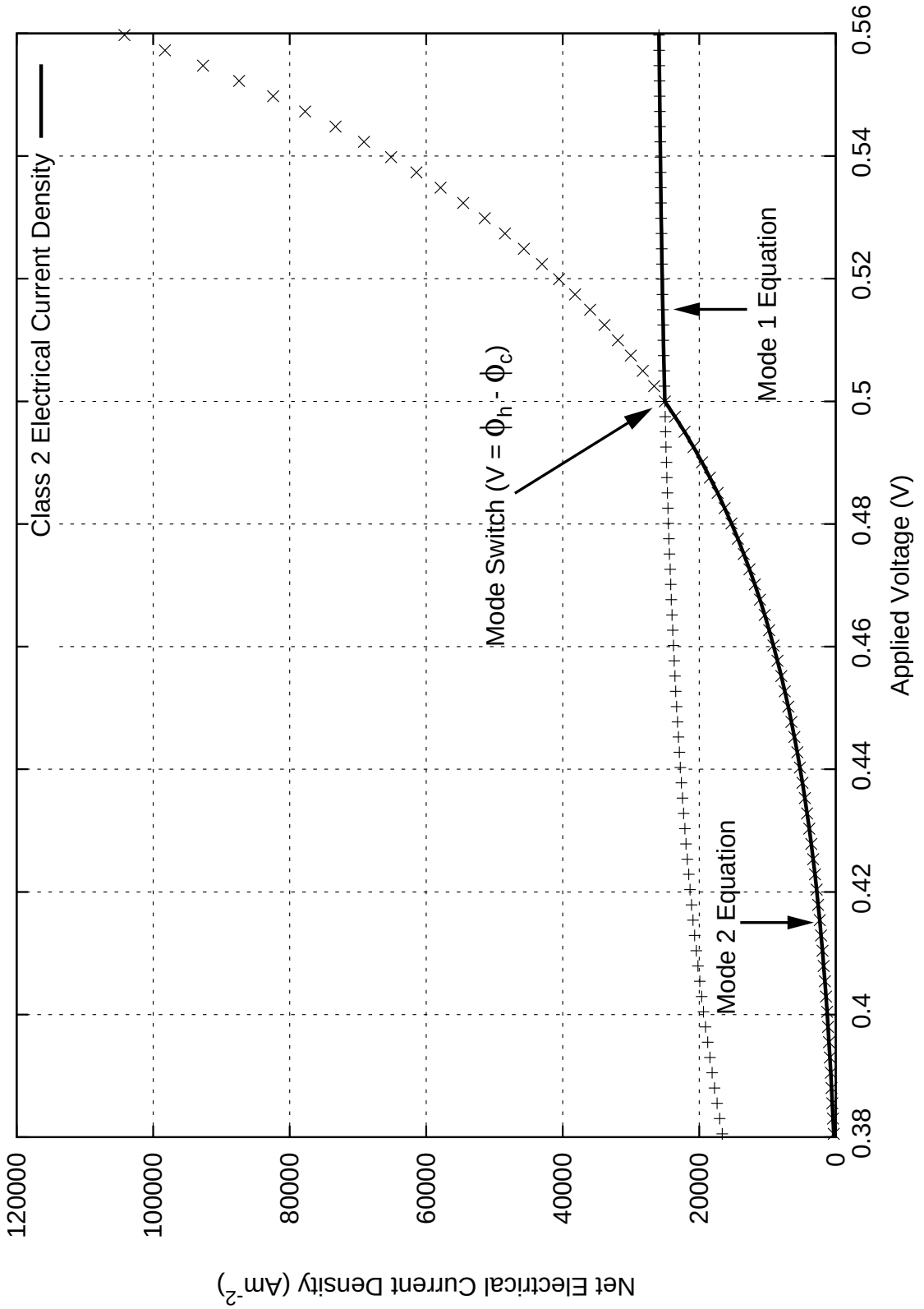


Figure 3.7: Electrical Current Density of a Class 2 device showing Mode Switch at $V = \phi_h - \phi_c$

### 3.4.3 Class 2 Device Modes

This section will illustrate a Class 2 device as it goes from Mode 2 ($\phi_h > \phi_c + V$) to Mode 1 ($\phi_c + V > \phi_h$) at the critical bias $V = \phi_h - \phi_c$. In order to compare with the Class 1 device already used as an example the same device parameters as in section 3.2.1 are used: $\phi_c = 0.7$ eV, $T_c = 500$ K and $T_h = 700$ K. According to Eq. 3.22 the barrier height of the hot side, $\phi_h$, must be larger than 1.021 V. We set $\phi_h = 1.2$ eV. This value is somewhat arbitrary but causes the switch in operating modes to be seen clearly in Figs. 3.7, 3.8 and 3.9. It should be noted that the Mode 1 curve on each of the figures is the curve that would be obtained for a Class 1 device ($\phi_c \geq \phi_h$) with $\phi_c = 0.7$ eV. These lines can be used as a direct comparison between the performance of a Class 1 and Class 2 device.

The net electrical current for a Class 2 device is shown in Fig. 3.7. The device switches modes at $V = \phi_h - \phi_c = 1.2$ V $- 0.7$ V $= 0.5$ V. The bias at which the Mode 2 electrical current becomes positive, according to Eq. 3.20, is 0.372 V but is not shown on the plot. The reason for this is that the same code was used to generate each of the figures. As Eq. 3.21 states, the bias at which the net heat current becomes positive is larger than 0.372 V. In order to avoid a numerical singularity when calculating the efficiency the minimum bias was set at a value higher than $V^{Jq}_{0\ (Class\ 2)}$. Numerically it was found that $V^{Jq}_{0\ (Class\ 2)} < 0.38$ V for this device. The bias at which the Mode 1 electrical current becomes positive is still 0.323 V as in Fig. 3.2.

The net heat current for a Class 2 device is shown in Fig. 3.8. As is expected, the bias where the Mode 2 equation becomes positive is again less than that for the Mode 1 equation. The efficiency of the Class 2 device is shown in Fig. 3.9. Because of the mode switch for Class 2 devices, the correct lines for these devices are given by the solid line in each of the figures.

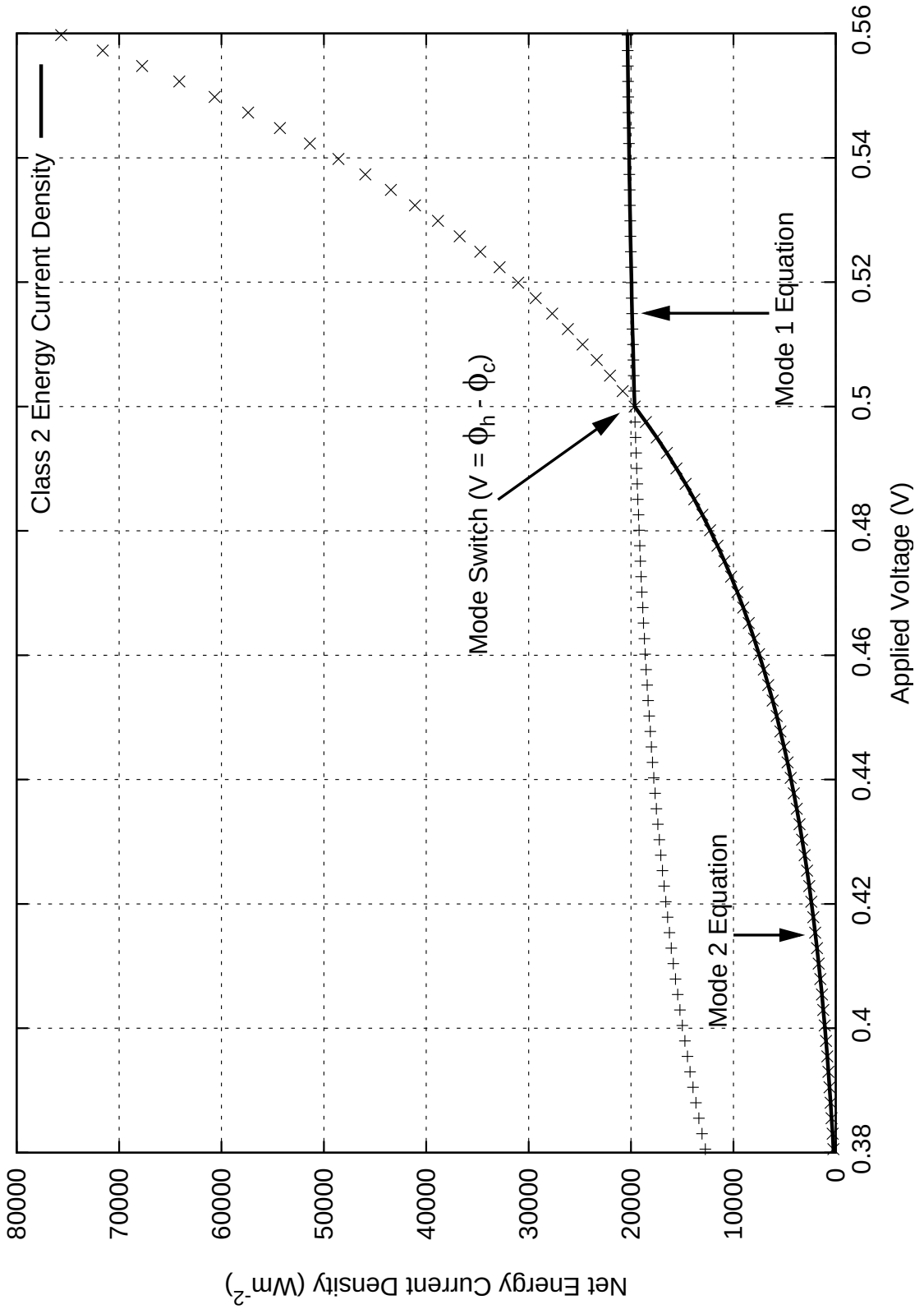


Figure 3.8: Heat Current Density of a Class 2 device showing Mode Switch at $V = \phi_h - \phi_c$

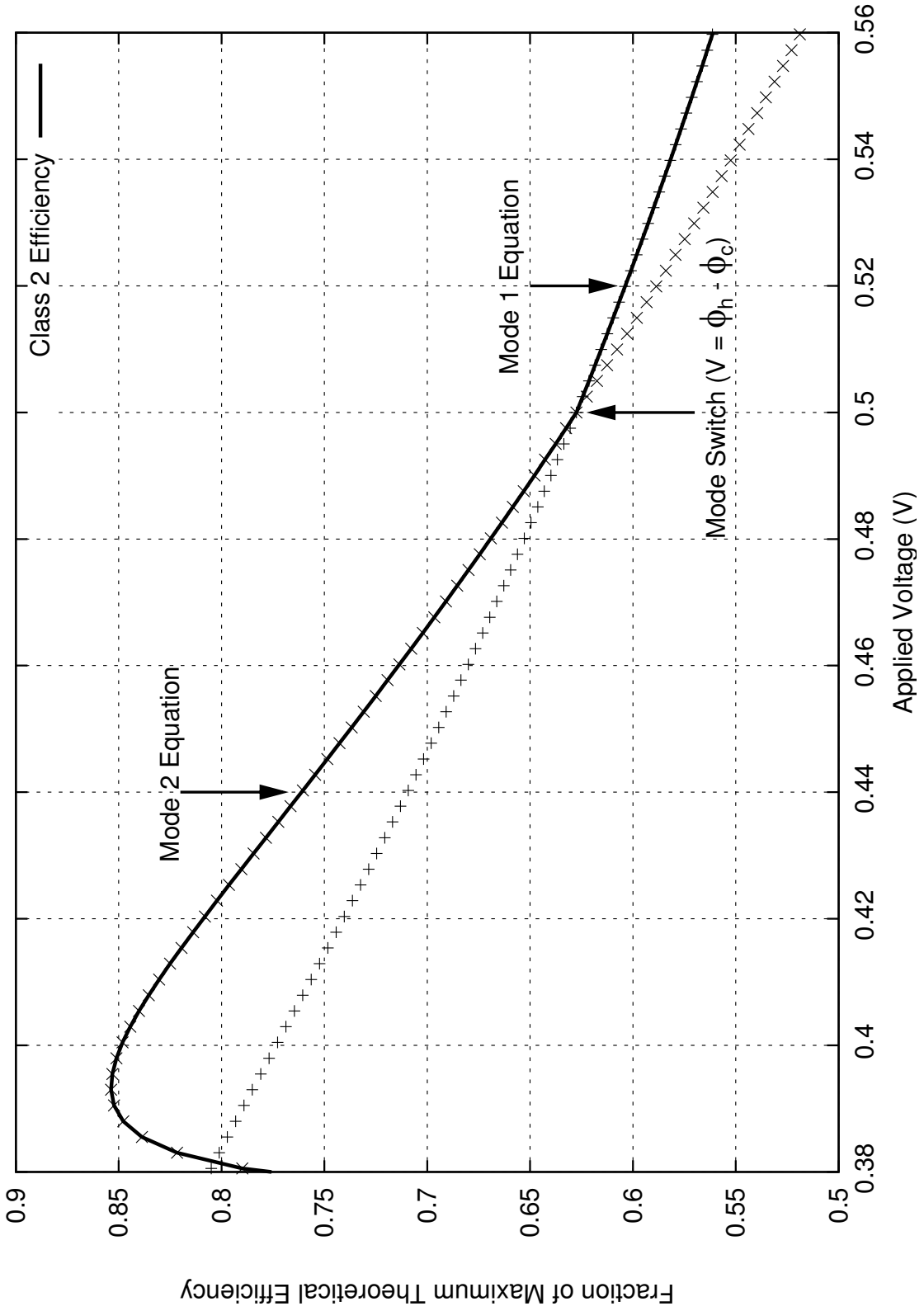


Figure 3.9: Efficiency of a Class 2 device showing Mode Switch at $V = \phi_h - \phi_c$

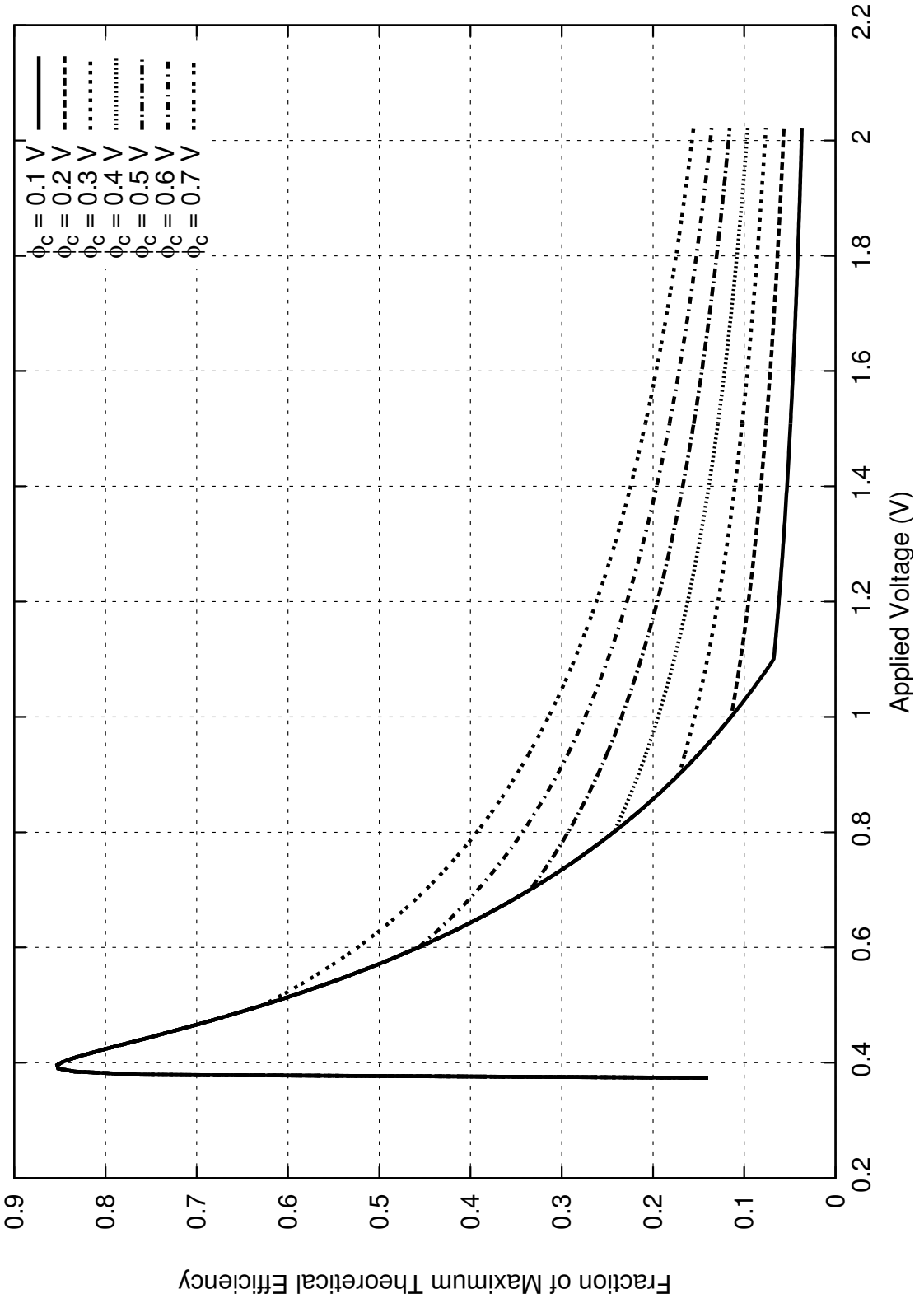


Figure 3.10: Efficiency of a Class 2 device with $\phi_c$ varying from 0.1 V to 0.7 V

### 3.4.4 Effect of $\phi_c$ on Class 2 Device Operation

By rearranging Eq. 3.22 we can obtain the maximum $\phi_c$ allowable for a given $T_c$, $T_h$ and $\phi_h$ that keeps the Class 2 device operating in Mode 2. For the device being considered this maximum value is $\phi_{c\,(max)} = 0.82$ V. Fig. 3.10 shows the same Class 2 device as above but this time with the cold barrier, $\phi_c$, varying from 0.1 V – 0.7 V in 0.1 V increments. While remaining in Mode 2, each device has the same maximum efficiency occurring at the same bias. This is to be expected since $\phi_c$ does not appear in the equations for this mode. Therefore, the only difference between each device is the critical bias ($V_{crit} = \phi_h - \phi_c$) at which they switch from Mode 2 to Mode 1. After this switch they then operate exactly the same as a Class 1 device with the same cold electrode barrier height.

### 3.4.5 Comparison between Class 1 and Class 2 Devices

As already stated, the bias at which Class 1 electrical and heat currents become positive is less than that for Class 2 devices. For a bias less than the critical voltage both the electrical and energy currents are larger for Class 1 devices meaning that more energy is removed for Class 1 devices. After the mode switch electrical and heat currents, and therefore efficiencies, are equal for both classes.

Fig. 3.11 shows the efficiency of both class of device versus applied bias. Both devices have a cold electrode with a barrier height of 0.7 eV. The Class 2 device has a hot electrode with a barrier height of 1.2 eV. In this figure the Class 2 device has a slightly higher efficiency but by referring to Fig. 3.8 it can be seen that the Class 1[1] device removes a substantial amount more heat from the cold electrode. By sacrificing a small reduction in efficiency a large increase a cooling power can be achieved.

Fig. 3.12 shows the maximum efficiency versus barrier height for both classes of device.

[1] remembering that the Mode 1 equations give the currents for the equivalent Class 1 device with the same $\phi_c$.

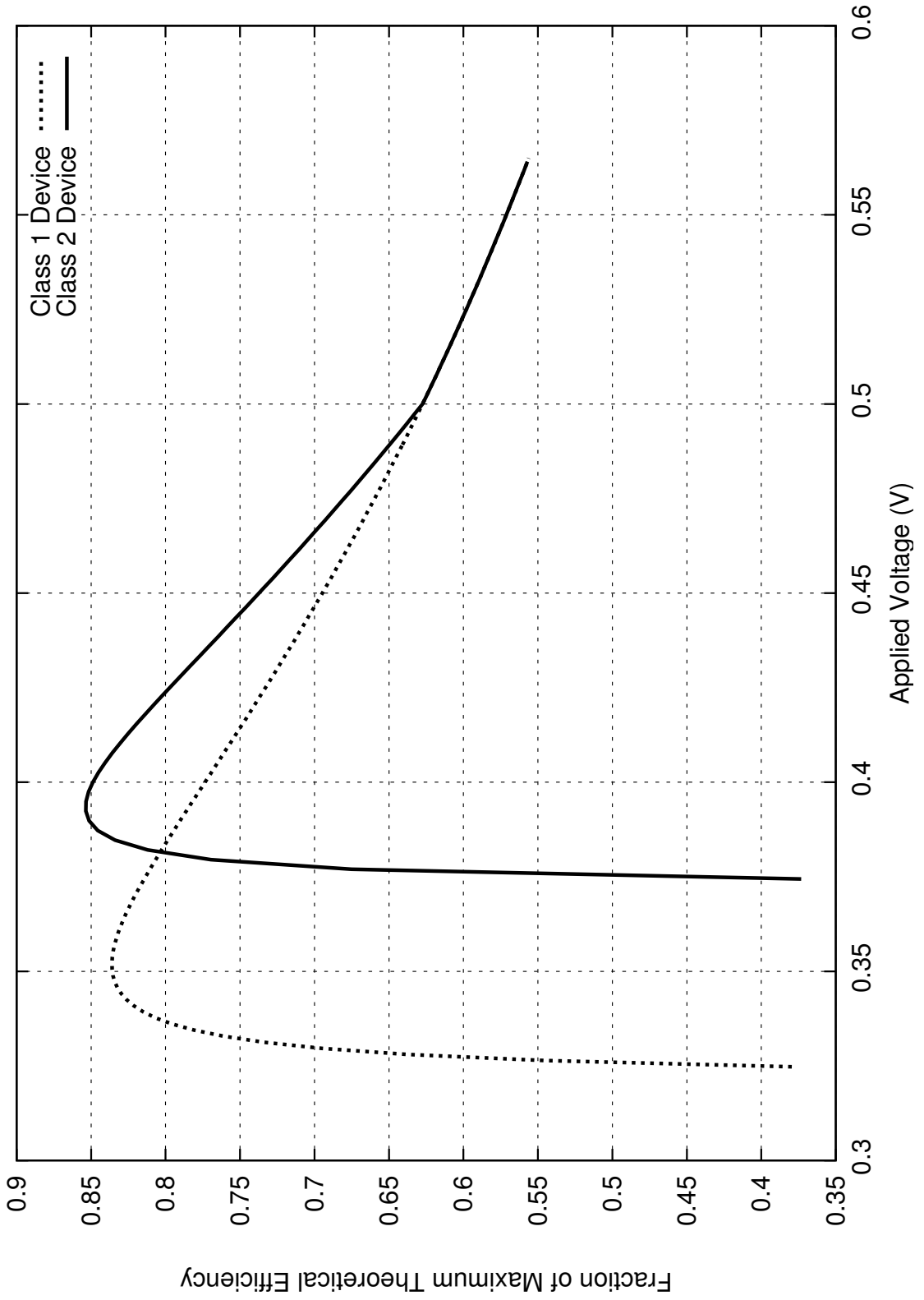


Figure 3.11: Efficiency of Class 1 and Class 2 devices with $\phi_c = 0.7$ eV and $\phi_h = 1.2$ eV (applicable to Class 2 Device only).

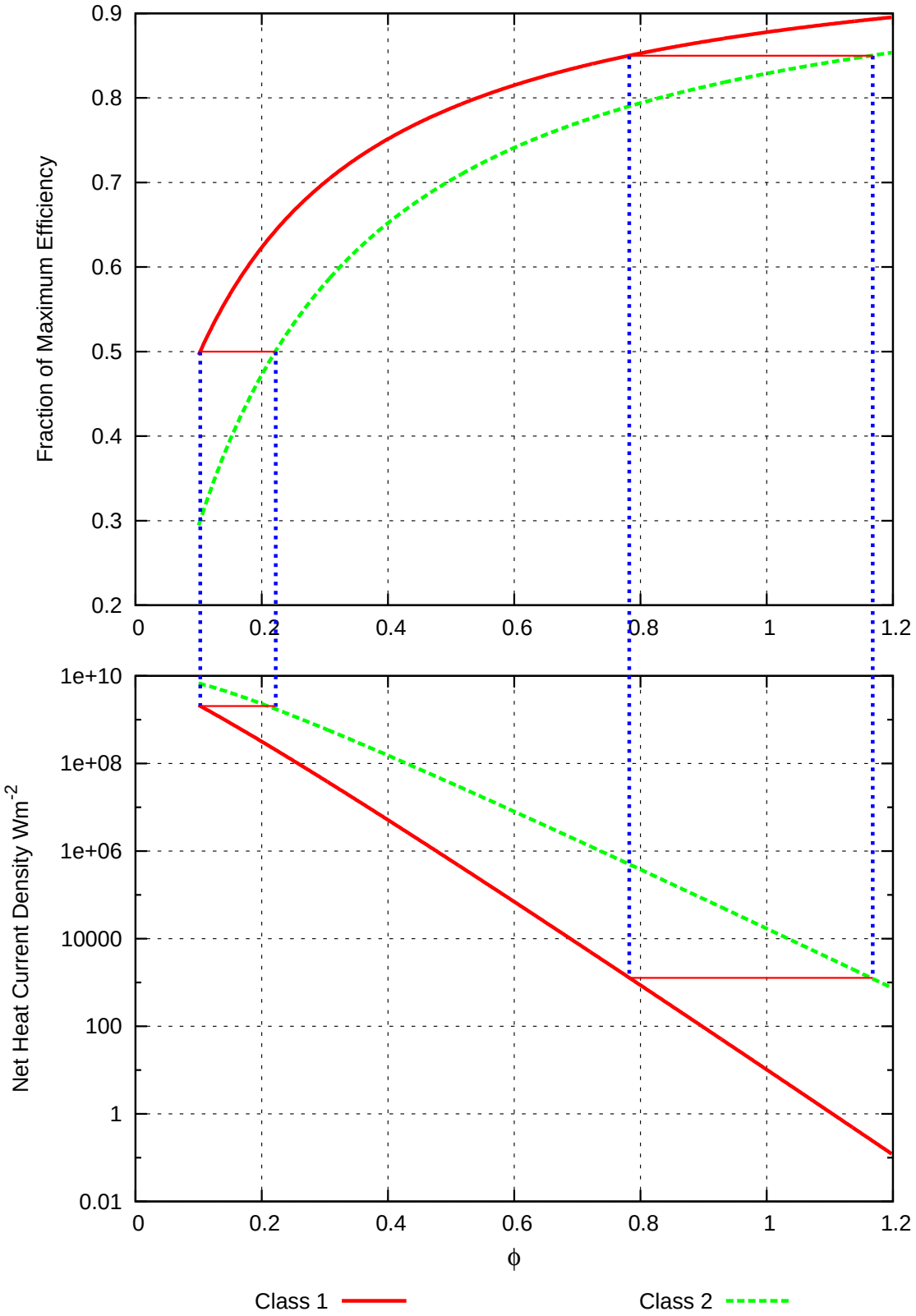


Figure 3.12: Comparison of performance between Class 1 ($\phi = \phi_c$) and Class 2 devices ($\phi = \phi_h$). Both devices have a $T_h = 700$ K and $T_c = 500$ K.

For Class 1 devices the barrier being varied is on the cold side. For Class 2 devices the barrier being varied is on the hot side. For the Class 2 device the cold barrier height is set small enough that the device is always operating in Mode 2 (as section 3.4.4 discussed the only effect the cold side barrier height has on Class 2 devices is the bias at which the device switches from Mode 2 to Mode 1. By setting $\phi_c$ small enough Mode 2 operation can be ensured. This is done because Class 1 devices always operate in Mode 1 and we want a direct comparison between Class 1 and Class 2 devices).

For a given barrier height the applied bias is varied until the maximum efficiency possible for that barrier height is found. This optimal applied bias is then used to calculate the heat current being removed from the cold electrode. By inspecting Fig. 3.12 it is seen that the thermal efficiency approaches the maximum theoretical efficiency as the barrier height increases. It should be noted however that this is the case only when the thermal resistance is considered infinite. If a finite thermal resistance is used then as the barrier height is increased, $J_{qc}^{sat}(= J_{qc}^{c})$ reduces until it is less than the heat current due to conduction and the device can never operate as a refrigerator. This will be discussed further in section 3.5.2. As the barrier height becomes larger the amount of heat removed from the cold side reduces substantially, in agreement with the statement made earlier.

Fig. 3.13 gives an expanded view of the heat current being removed from the cold side at two points where the efficiency is the same for both devices. As can be seen the barrier height at which the two classes have the same efficiency is different for the two classes. For example Class 1 devices have an efficiency that is 50% of the maximum when the barrier height is 0.103 eV. Class 2 devices have the same efficiency at a barrier height of 0.222 eV. By comparing the heat current being removed at these barrier heights with an applied bias that maximises the efficiency at these points it can be seen that Class 1 devices remove more heat from the cold side, thus giving better overall device performance.

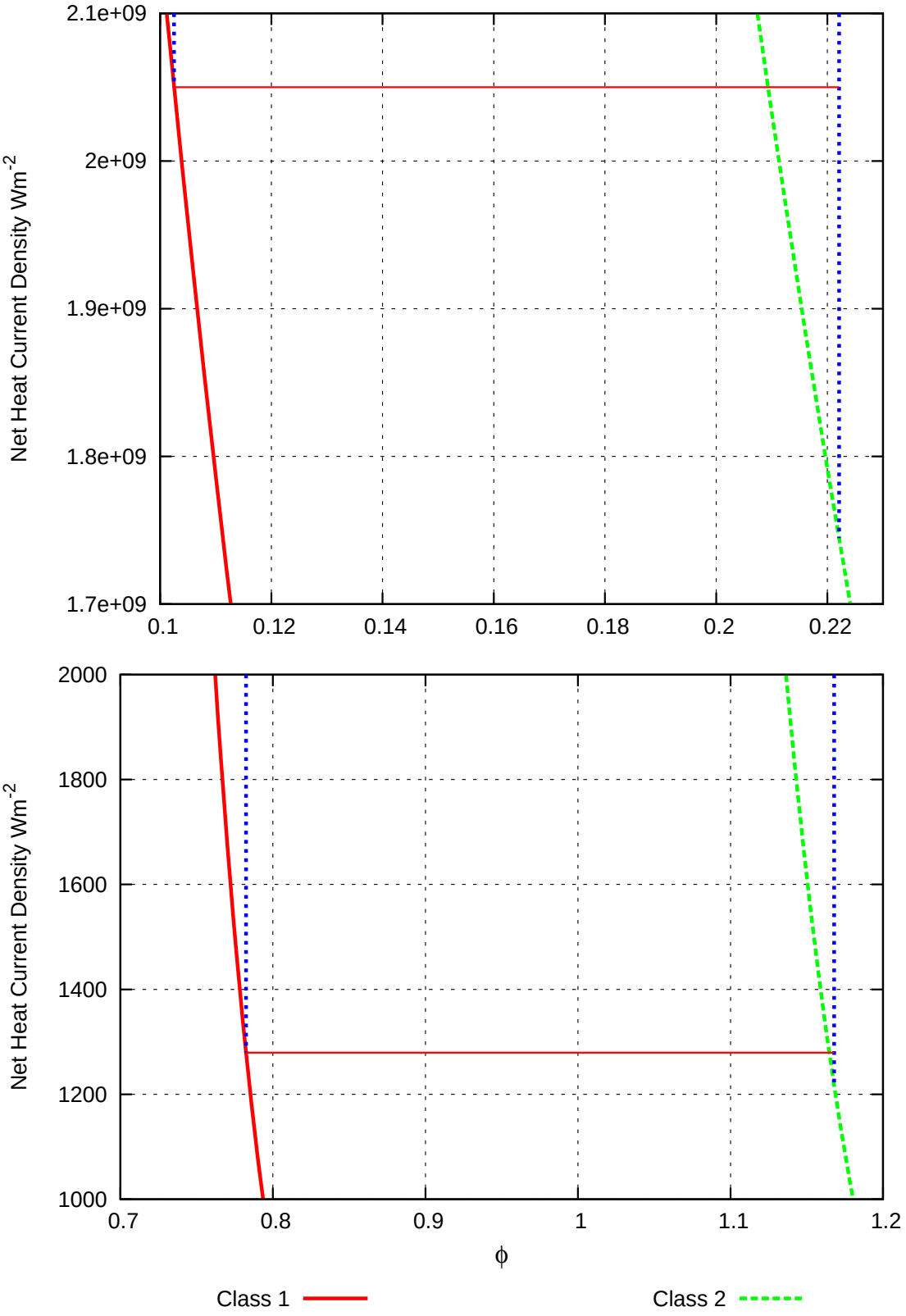


Figure 3.13: Comparison of performance between Class 1 ($\phi = \phi_c$) and Class 2 devices ($\phi = \phi_h$). Both devices have a $T_h = 700$ K and $T_c = 500$ K.

### 3.4.6 Conclusion – Class 1 Devices perform better

This section has shown that there are two different classes of single-barrier device that need to be considered. Class 1 devices have $\phi_c \geq \phi_h$ while Class 2 devices have $\phi_h > \phi_c$. The last section has shown that for a given efficiency Class 1 devices remove more heat from the cold side than Class 2 devices. In addition to this, solving equations for Class 2 devices is more complicated due to the devices switching from Mode 2 to Mode 1. Also, there is the added complexity of having to solve $V_{0\ (Class2)}^{Jq}$ (Eq. 3.21) numerically. For these reasons, only Class 1 devices will be considered further. By considering only Class 1 devices, $\phi_h$ does not enter into the transport equations as long as the devices are operating in Mode 1 ($V \geq 0$). This condition will be satisfied whenever any Class 1 device is operating as a refrigerator.

## 3.5 Consideration of Finite Thermal Resistance

In the preceding sections the thermal resistance of the devices has been considered infinite. This assumption was useful for comparing Class 1 and Class 2 devices, but in a real device the finite thermal resistance must not be neglected. As discussed by Mahan et al. [3] and by referring to Eq. 3.13, the net heat current removed from a surface is reduced by heat conduction represented by the term $\Delta T/R^{th}$. The thermal resistance, $R^{th}$, is directly proportional to the width, $L$, of a given barrier. To increase $R^{th}$ we must increase $L$ accordingly. But, in order for our assumptions of section 3.1.1 to remain valid, (namely that carriers move ballistically through the barrier), the barrier width must be less than the *mean free path* for the carriers in the barrier material. Mahan et al. [3] mention that there is little data available on the mean free path of carriers when travelling perpendicular to the layers in a heterostructure. They do note that there is quite a substantial amount of data for motion along the layers. They reason that if they are similar, then the mean free path for electrons in the barriers is 50 – 100 nm.

### 3.5.1 Mobility, Mean Free Path of Carriers, Thermal Resistivity and Thermal Resistance

The mean free path in a bulk material can be calculated by starting with the definition of carrier mobility [36, 37]:

$$\mu = \frac{q\langle\tau_m\rangle}{m} \tag{3.23}$$

where $q$ is the charge and $m$ is the mass of the carrier. $\langle\tau_m\rangle$ is defined as the *momentum relaxation time* and is the average time between collisions of the carrier. The mobility is a readily available quantity for many materials, so from this the momentum relaxation time can be calculated:

$$\langle\tau_m\rangle = \frac{\mu m}{q} \tag{3.24}$$

For an ideal gas, which can be extended to include carriers in the free-electron model, the mean velocity of carriers is given by [37]:

$$\langle v_x^2\rangle = \langle v_y^2\rangle = \langle v_z^2\rangle = \frac{k_B T}{m^*} \tag{3.25}$$

So the mean distance a carrier travels before experiencing a collision (*mean free path*) is given, in metres, by:

$$\text{MFP} = \langle v_x\rangle\langle\tau_m\rangle = \frac{\mu}{q}\sqrt{k_B T m^*} \tag{3.26}$$

As an example, using available values for electrons in bulk Gallium Arsenide [38] at 300 K, $\mu$ = 8500 cm$^2$/V·s = 0.85 m$^2$/V·s and $m$ = 0.063$m_e$, gives MFP$_{GaAs} \approx$ 82 nm. This value is in agreement with the range quoted above. These values used to calculate the MFP above are for bulk GaAs so do not give any information about electron motion either along or perpendicular to the heterostructure layers. Although some very high mobility materials ($>10^7$ cm$^2$/V·s) [39] have been fabricated using modulation doped $Al_xGa_{1-x}As$/GaAs heterostructures, many of these values are achieved at low temperature. The devices of interest here are to operate near room temperature. At these higher temperatures ($\sim$300 K) the mobilities converge to a much lower value of between 5000 – 10000 cm$^2$/V·s for GaAs. For GaAs at least, we can then assume that the barrier widths must be in the range quoted by Mahan.

As a comparison, electrons in bulk InSb have a mobility of up to $7.7\times10^4$ cm$^2$/V·s [40] and effective mass of $m$ = 0.014$m_e$. This corresponds to a mean free path of MFP$_{InSb} \approx$ 349 nm. A very large mean free path means that the barrier regions can be much larger. A larger barrier generally means a larger thermal resistance, depending on the thermal resistivity of the material.

One of the design goals will be to find a material with a large carrier mobility and also a good thermal resistivity. Usually, semiconductors with small band gaps have small effective masses due to a large curvature of $E$ vs. $\boldsymbol{k}$ caused by repulsion of the eigenstates [41]. The small effective mass results in a high mobility due to the inverse relationship given in Eq. 3.23. As will be seen in later sections, another design criterion will be to have a large effective mass of carriers. This term appears in the numerator of Richardson's constant and a larger value results in a larger efficiency and/or power output, depending on the applied bias.

To give a practical example of maximising thermal resistance, the thermal resistivity of GaAs is 1.82 cm·K/W [42]. The maximum thermal resistance per unit area for GaAs, under conditions for ballistic transport, is then:

$$\begin{aligned} R^{th}_{GaAs} &= 1.82 \times 10^{-2}\ \text{m·K/W} \times 82 \times 10^{-9}\ \text{m} \\ &= 1.49 \times 10^{-9}\ \text{m}^2\text{·K/W} \end{aligned} \tag{3.27}$$

For InSb, the thermal resistivity is 5.56 cm·K/W [40]. The maximum thermal resistance per unit area for InSb is then:

$$\begin{aligned} R^{th}_{InSb} &= 5.56 \times 10^{-2}\ \text{m·K/W} \times 349 \times 10^{-9}\ \text{m} \\ &= 19.42 \times 10^{-9}\ \text{m}^2\text{·K/W} \end{aligned} \tag{3.28}$$

This is a very important result because, as will been shown in the next section, the thermal resistance of the barrier material is one of the most important factors in device performance. Section 4.6.2 will compare the maximum cooling possible for 10-barrier structures using different semiconductor materials.

### 3.5.2 Effect of Thermal Resistance on Device Performance

Fig. 3.14 shows the efficiency and net electrical and energy currents removed from the cold side of a Class 1 single-barrier thermionic device with a varying finite thermal resistance. The thermal resistance ranges from the value for InSb given in Eq. 3.28 to 10 and 100 times this value. A temperature gradient of 20 K is used with $T_h$ set near room temperature ($300$ K). The barrier height of the devices ranges up to a value of 200 meV.

A realistic value for the thermal resistance (Eq. 3.28) reduces the thermal efficiency to less than 10% of the theoretical maximum at the two specified temperatures. As discussed in section 3.4.5, when a finite thermal resistance is considered the device efficiency does not increase asymptotically with an increase in barrier height. The larger barrier height reduces the saturation heat current removed from the cold side. Even when a very large bias is applied, less heat is removed from the cold electrode thermionically than is deposited due to heat conduction. Devices with large potential barriers, then, can never operate as a refrigerator.

For a temperature gradient of 20 K (a realistic requirement for domestic refrigeration at room temperature) we see that this induces a large conduction heat current per unit area in GaAs of

$$J_q^{cond} = \frac{20 \text{ K}}{1.49 \times 10^{-9} \text{ m}^2\cdot\text{K/W}} = 1.34 \times 10^{10} \text{ Wm}^{-2} \tag{3.29}$$

For a single-barrier device made of GaAs with $\phi_c = 70$ meV, $T_c = 280$ K and $T_h = 300$ K, the *saturation heat current* [2] , $J_{qc}^{sat}$, leaving the cold side is 6.11$\times 10^8$ Wm$^{-2}$. This is 20 times less than the conduction heat current calculated above. Even when an unrealistic device with $\phi_c = 0$ eV is considered, the saturation heat current removed from the cold side is only 4.54$\times 10^9$ Wm$^{-2}$. This means that for a single-barrier device made from a material like GaAs, even when operating under saturation conditions, it is impossible to get 20 K cooling at room temperature. This amount of cooling is necessary for

[2] when the applied bias is so large that that $J_{qc}^{h} \to 0$ and $J_{qc}^{out} \to J_{qc}^{sat} = J_{qc}^{c}$ in Eq. 3.10

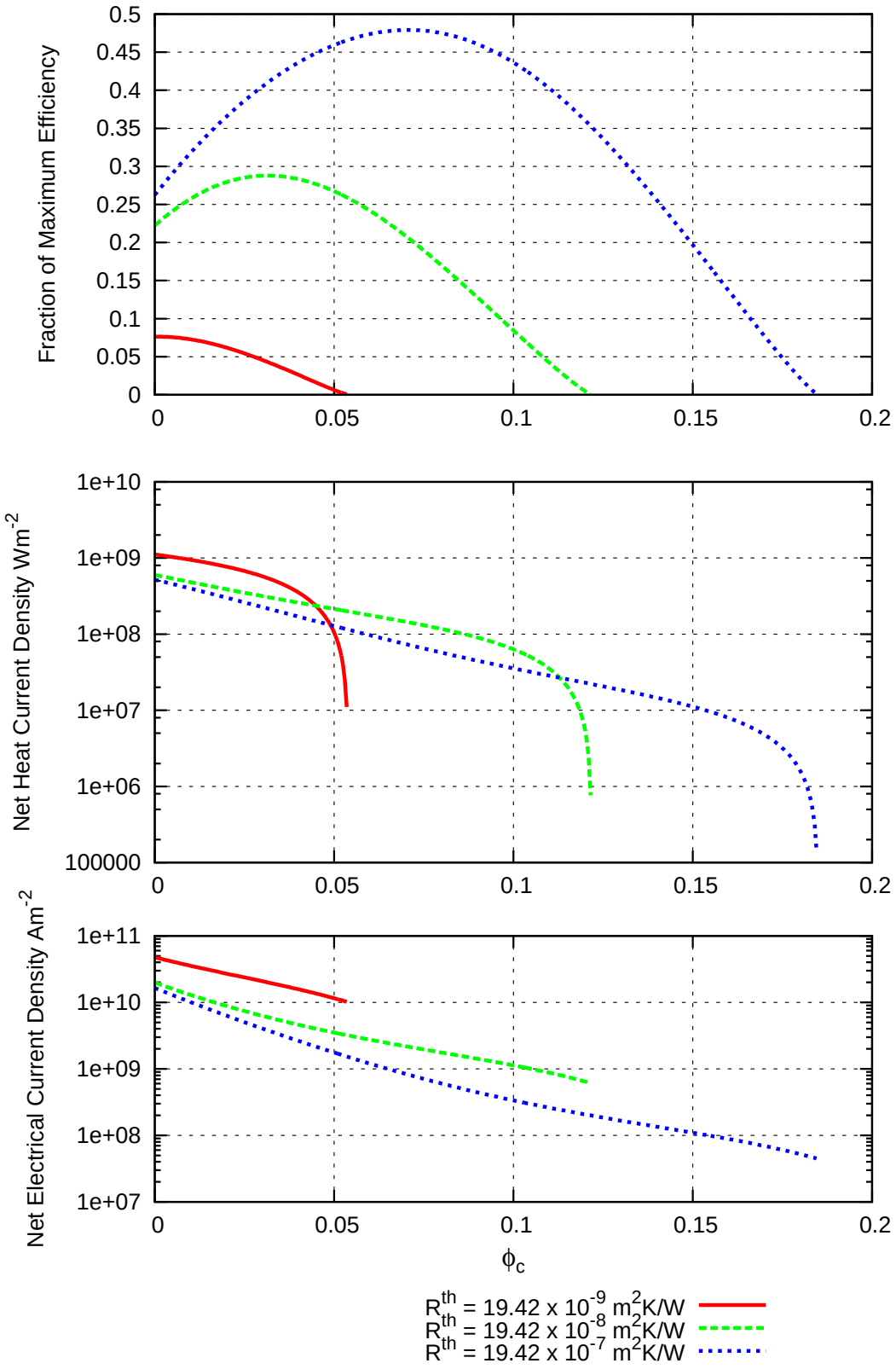


Figure 3.14: Performance of a Class 1 device with varying Thermal Resistance ($T_c = 280$ K, $T_h = 300$ K).

domestic refrigeration. For other applications, such as integrated thermionic cooling for temperature stabilisation in a laser, this large temperature difference may not be necessary. Regardless, a large amount of power is needed to just neutralise this large back-flow of heat current, let alone actually induce any cooling in the device. As a result no curve for GaAs is shown on Fig. 3.14 because there is no barrier height that will cause the device to operate as a refrigerator at these temperatures.

### 3.6 Effect of Temperature on Device Performance

Fig. 3.15 shows the efficiency and net electrical and energy currents removed from the cold side of a Class 1 single-barrier thermionic device with a varying temperature, $T_h$, of the hot electrode. The thermal resistance is equal to value for InSb given in Eq. 3.28 ($19.42 \times 10^{-9}$ m$^2$·K/W). A temperature gradient of 20 K is used with $T_h$ set at $300$ K, $500$ K and $700$ K.

By increasing the operating temperature of the device and keeping the temperature gradient across the device constant it can be seen from the figure that the devices perform better at elevated temperatures — that is, the efficiency and heat removed from the cold side both increase. This property is consistent with the maximum theoretical thermal efficiency of reversible refrigerators given by Eq. 3.17. For a fixed temperature gradient, the efficiency increases proportionally with $T_c$. This property of the devices may be of use in other applications. By reversing the operating mode the devices could be used for power generation at elevated temperatures.

### 3.7 Handling Negative Applied Bias for Class 1 Devices

Because Eq. 3.11 states that the applied bias must have a minimum positive value in order for the devices to be operating as a refrigerator, negative applied biases may appear to be of little interest. But, a negative applied bias needs to be considered when generating $I$-$V$ curves for the devices. Referring to Fig. 3.16, special care must be taken whenever $\phi_R - V_A > \phi_L$ for a given barrier. Under these conditions Eq. 3.5, Eq. 3.13 and Eq. 3.14

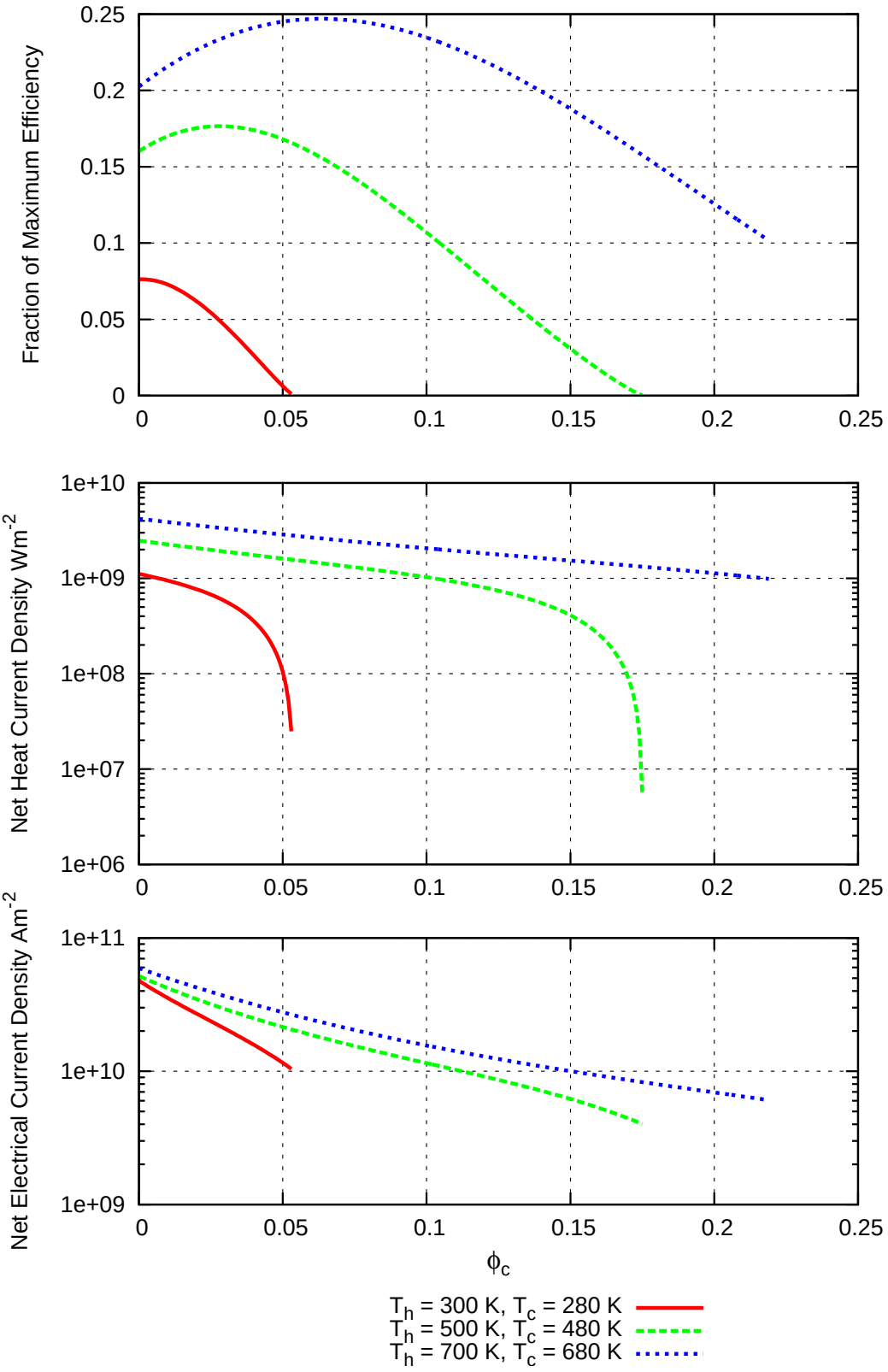


Figure 3.15: Performance of a Class 1 device with varying Hot Temperature ($R^{th} = 19.42 \times 10^{-9}$ m$^2$·K/W).

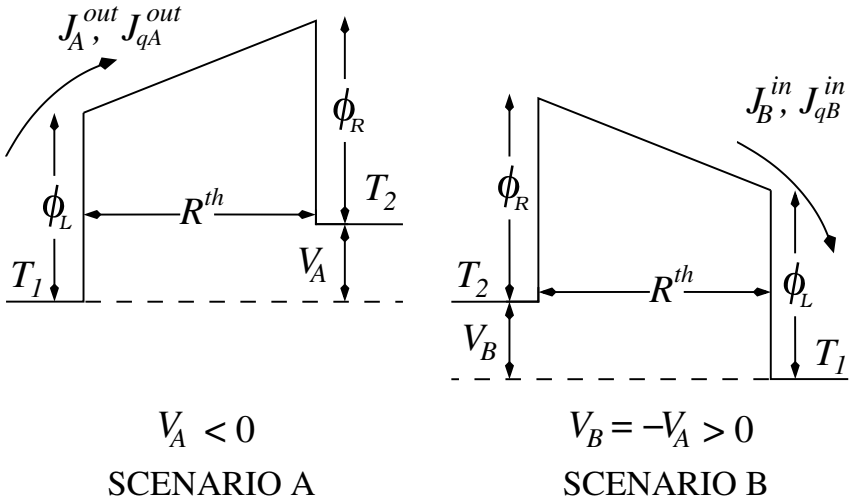


Figure 3.16: Equivalence of Scenario A, with negative applied bias and Scenario B, with positive applied bias. In both cases $\phi_R + |V_{A,B}| > \phi_L$.

are no longer valid. Under Scenario A, carriers must overcome $\phi_R - V_A$ in order contribute to current flow, as opposed to the usual case when carriers must overcome $\phi_L + V = \phi_c + V$ as in Fig. 3.1. In order for the system of equations derived so far to remain valid, the equivalence of Scenario A and Scenario B can be used. If we assume Class 1 devices we can set $\phi = \phi_L = \phi_R$

For Scenario A, the net electrical and heat currents leaving the left electrode are given by:

$$J_A^{out} = AT_1^2 \exp\left(-\frac{q\left(\phi - V_A\right)}{k_B T_1}\right) - AT_2^2 \exp\left(-\frac{q\phi}{k_B T_2}\right) \tag{3.30}$$

and

$$\begin{aligned} J_{qA}^{out} &= \left(\phi - V_A + \frac{2k_B T_1}{q}\right) AT_1^2 \exp\left(-\frac{q\left(\phi - V_A\right)}{k_B T_1}\right) \\ &- \left(\phi - V_A + \frac{2k_B T_2}{q}\right) AT_2^2 \exp\left(-\frac{q\phi}{k_B T_2}\right) \\ &- \left(\frac{T_2 - T_1}{R^{th}}\right) \end{aligned} \tag{3.31}$$

By simply switching the temperatures of the two electrodes, taking the negative of the applied bias and keeping $\phi$ constant we achieve the equivalent device given in Scenario

B. The net electrical and heat currents then entering the right electrode are given by:

$$J_B^{in} = AT_2^2 \exp\left(-\frac{q\phi}{k_B T_2}\right) - AT_1^2 \exp\left(-\frac{q\left(\phi + V_B\right)}{k_B T_1}\right) \tag{3.32}$$

and

$$\begin{aligned} J_{qB}^{in} &= \left(\phi + V_B + \frac{2k_B T_2}{q}\right) AT_2^2 \exp\left(-\frac{q\phi}{k_B T_2}\right) \\ &- \left(\phi + V_B + \frac{2k_B T_1}{q}\right) AT_1^2 \exp\left(-\frac{q\left(\phi + V_B\right)}{k_B T_1}\right) \\ &- \frac{T_1 - T_2}{R^{th}} \end{aligned} \tag{3.33}$$

which are the already familiar forms given by Eq. 3.5 and Eq. 3.14.

By remembering that $V_B = -V_A$ we have the equivalence that:

$$J_A^{out} = -J_B^{in} \tag{3.34}$$

$$J_{qA}^{out} = -J_{qB}^{in} \tag{3.35}$$

so that Eq. 3.5, Eq. 3.13 and Eq. 3.14 can still be used by simply switching the temperatures of both sides and making the applied bias positive.

One interesting point to note is that when a negative bias is applied the heat current no longer saturates because of the appearance of $V$ in the linear part of the heat current equations. As a larger bias is applied more heat is deposited in the electrode equal to $VJ$ so the heat current increases linearly with increasing $V$.

### 3.8 Conclusion

This chapter has introduced single-barrier semiconductor thermionic devices and two classes of device were discussed. Class 1 devices can only ever operate in Mode 1, whilst Class 2 devices can operate in both Mode 1 and Mode 2, depending on the applied bias. It was found that Class 1 devices always have better performance than Class 2 devices. In addition to this, Class 1 devices are much easier to model than Class 2 devices. As a result only Class 1 devices need to be considered further. By assuming that $\phi_c \geq \phi_h$, Class 1 devices are ensured and $\phi_h$ no longer appears in device equations.

The comparison of Class 1 and Class 2 intentionally neglected heat currents due to conduction, which always works to reduce device performance. The reasoning behind this is that the conduction heat current is a constant term (under the assumptions of the simple model used) which is the same in the equations for both class of device.

In reality the parasitic heat current due to heat conduction can not be ignored in a real device. When the finite thermal resistance of two materials, GaAs and InSb, is included in the equations it is seen that this drastically reduces overall device performance. Indeed, no single-barrier thermionic device made of GaAs (assuming bulk material properties) can operate as a refrigerator at room temperature with a temperature gradient of 20 K. Even for a device made from a material with 'good' thermal resistance, like InSb, a maximum efficiency of less than 10% of the maximum theoretical efficiency is possible.

The next chapter will investigate multiple barrier semiconductor devices and it will be shown that these perform much better.

# CHAPTER 4

# THERMIONIC REFRIGERATION IN A MULTIPLE-BARRIER SEMICONDUCTOR DEVICE

## 4.1 Motivation for using a Multilayer System

Chapter 3 showed that the performance of semiconductor thermionic devices is dramatically reduced by the conduction heat current, $\Delta T/R^{th}$, appearing in the heat current equations. Two methods of reducing this term are either i) increasing $R^{th}$ or ii) reducing $\Delta T$. As discussed in section 3.5.1, even with careful material selection, $R^{th}$ can only be increased by a limited amount, dependent on the mean free path of carriers in the material. Even for InSb, the thermal resistance (per unit area) is still very small because of the device geometries required for ballistic transport. The other option is to reduce the temperature gradient across the barrier region of the devices.

### 4.1.1 Decreased temperature gradient across barriers

Mahan et al. [3] proposed using a multi-barrier system by reasoning that using $N$ barriers identical to a single barrier would reduce the temperature gradient across each barrier to $\sim \Delta T/N$; effectively reducing the back-flow by a factor of $N$. This chapter will investigate this proposal and show that increased heat power removal is possible with a multi-barrier system.

### 4.1.2 Reduced thermal conductivity of superlattices

One other, perhaps more important, advantage of using a multi-barrier system stems from experimental work carried out on the thermal conduction of superlattices as compared to bulk materials [17]. Following this experimental work, theoretical work was carried out

to explain the reduced thermal conductivity of superlattice as compared to bulk materials. By solving the Boltzmann Transport Equation for phonons and using various phonon interface scattering mechanisms, Chen [43] successfully modelled the majority of these experimental results. The theoretical results presented show that the thermal conductivity of the superlattices is roughly proportional to the square root of the thickness of each of the layers. Because the total width of the superlattices is held constant this is equivalent to saying that, for a fixed sample size, the thermal conductivity decreases proportionally to the number of barriers in the sample. Section 4.6.2 will investigate the maximum cooling for 10-barrier structures using different materials. As a first approximation this result will be considered by simply multiplying the thermal resistivity of bulk materials by the square root of the number of barriers used (10). The next section will introduce development of a numerically-based system to solve the electrical and energy balance equations for multiple-barrier thermionic devices.

## 4.2 Numerical Methods

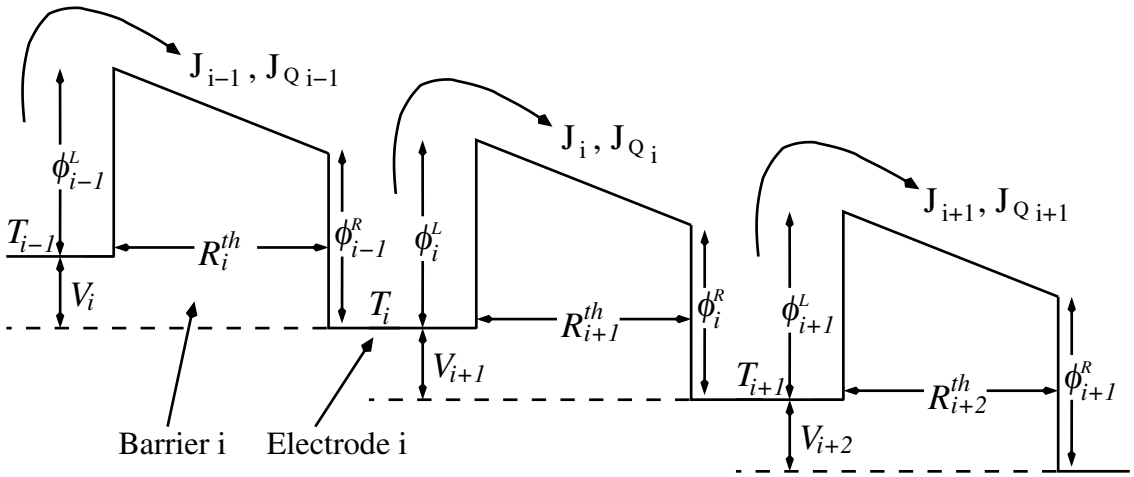


Figure 4.1: Schematic of Generic Multilayer Thermionic Cooler

Building on the system developed in Chapter 3, it is an easy task conceptually to extend the system to handle multiple barrier structures. That is, we consider a number of single-barrier devices sandwiched together as shown in Fig. 4.1. In this figure $\phi_i^L$ may or may

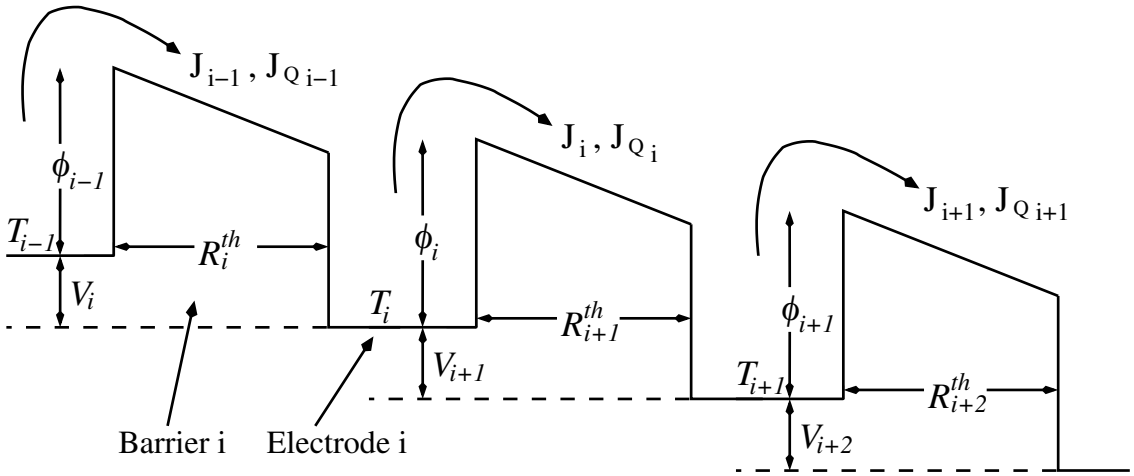


Figure 4.2: Multilayer Thermionic Cooler using only Class 1 devices

not be larger than $\phi_i^R$. By taking the conclusions of Chapter 3 into consideration, that Class 1 devices perform better than Class 2 devices, we need only consider the system shown in Fig. 4.2. By considering only Class 1 single-barrier devices ($\phi_i^L \geq \phi_i^R$) we can ignore $\phi_i^R$ altogether and use:

$$\phi_i^L \rightarrow \phi_i \tag{4.1}$$

Dealing with such a system analytically becomes cumbersome, although it can be done quite effectively [13]. One approach to solving the system equations is to use numerical methods. It has a number of distinct advantages such as new physics being easily included into the model and the ability to read out any device parameters once convergence has been reached. In addition to this, solving the system under different conditions becomes easy once the basic system has been developed. For instance determining the bias at which a given device operates with maximum power output is easily found by solving the equations whilst varying the applied bias in the direction that increases the power output.

### 4.3 The System: Equations and Variables

A system has been developed to numerically model an 'ideal' $N$-barrier thermionic refrigerator. The term 'ideal' refers to a device in which the assumptions of section 3.1.1 are

upheld. Previous studies have included equations that are approximate solutions whose validity depends a number of assumptions. One such assumption made by Mahan, Sofo and Bartkowiak [3] is that the voltage across the device varies with respect to the temperature gradient across the device. In addition to verifying such assumptions, the model has shown value in being able to produce a profile of parameters varying across the device (such as the temperature and voltage of each electrode).

Fig. 4.2 shows the multilayer thermionic cooler that is to be simulated. The width of each barrier is less than the carrier mean free path and greater than the tunnelling width. In order to simulate the device, electrical and thermal continuity must be ensured at each electrode. To do so requires solving a number of simultaneous equations which are detailed below. In the proceeding sections, net electrical and energy currents are assumed to be positive when moving from left to right.

### 4.3.1 Electrical Currents

The *net* electrical current across the $i^{th}$ barrier (from left to right) is given by [44] (cf. Eq. 3.5):

$$J = J_i = AT_{i-1}^2 \exp\left(\frac{-q\phi_{i-1}}{k_B T_{i-1}}\right) - AT_i^2 \exp\left(\frac{-q\left(\phi_{i-1} + V_i\right)}{k_B T_i}\right) \quad i = 1, 2, \ldots, N \tag{4.2}$$

It should be noted that this current is constant throughout the device.

### 4.3.2 Heat Currents

Assuming net heat current flow is from left to right, the energy current leaving the $i^{th}$ electrode to the right is given as [44] (cf. Eq. 3.13):

$$\begin{aligned} J_{qi}^{out} &= \left(\phi_i + \frac{2k_B T_i}{q}\right) AT_i^2 \exp\left(\frac{-q\phi_i}{k_B T_i}\right) \\ &- \left(\phi_i + \frac{2k_B T_{i+1}}{q}\right) AT_{i+1}^2 \exp\left(\frac{-q(\phi_i + V_{i+1})}{k_B T_{i+1}}\right) \\ &- \frac{T_{i+1} - T_i}{R_{i+1}^{th}(N)} \qquad i = 1, 2, \ldots, N \end{aligned} \tag{4.3}$$

and the energy current entering the $i^{th}$ electrode from the left is given

as [44] (cf. Eq. 3.14):

$$\begin{aligned} J_{qi}^{in} &= \left(\phi_{i-1} + V_i + \frac{2k_B T_{i-1}}{q}\right) A T_{i-1}^2 \exp\left(\frac{-q\phi_{i-1}}{k_B T_{i-1}}\right) \\ &- \left(\phi_{i-1} + V_i + \frac{2k_B T_i}{q}\right) A T_i^2 \exp\left(\frac{-q\left(\phi_{i-1} + V_i\right)}{k_B T_i}\right) \\ &- \frac{T_i - T_{i-1}}{R_i^{th}(N)} \qquad\qquad i = 1, 2, \ldots, N \end{aligned} \tag{4.4}$$

where $R_i^{th}(N)$ is the thermal resistivity of the $i^{th}$ barrier for an $N$-barrier system. If increased thermal resistance due to interface scattering or superlattices is considered, $R_i^{th}(N) \neq R_i^{th}(N-1)$.

### 4.3.3 Continuity Equations

For continuity of heat current, $J_{qi}^{in} = J_{qi}^{out}$.

NOTE: $J_{qi}^{in} - J_{q(i-1)}^{out} = V_i J \neq 0$, where $V_i J$ is the work needed to produce cooling across the $i^{th}$ barrier.

For continuity of electrical currents $J_{i-1} = J_i = J_{i+1} \cdots$.

For an $N$-barrier system, there are $n = 2(N-1)$ unknowns and $n = 2(N-1)$ equations:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_{N-1} \\ x_N \\ \vdots \\ x_{2(N-1)} \end{bmatrix} = \begin{bmatrix} V_1 \\ \vdots \\ V_{N-1} \\ T_1 \\ \vdots \\ T_{N-1} \end{bmatrix} \tag{4.5}$$

$$\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} F_1 \\ \vdots \\ F_{N-1} \\ F_N \\ \vdots \\ F_{2(N-1)} \end{bmatrix} = \begin{bmatrix} J_1 - J_2 \\ \vdots \\ J_{N-1} - J_N \\ J_{q1}^{in} - J_{q1}^{out} \\ \vdots \\ J_{q(N-1)}^{in} - J_{q(N-1)}^{out} \end{bmatrix} \tag{4.6}$$

In this device configuration there are $N$ barriers — numbered $1$ to $N$; and $(N+1)$ electrodes — numbered $0$ to $N$.

The first $(N-1)$ variables in the vector, $\boldsymbol{x}$, are the biases across each barrier (or the bias of electrode $i$ with respect to the bias of electrode $i-1$). The bias across the last barrier is set and so is not an unknown. The second $(N-1)$ variables are the temperatures of each electrode numbered $1$ to $N-1$. The temperature of the $0^{th}$ and $N^{th}$ barriers are set and so are not unknowns.

The first $(N-1)$ equations are the electrical currents entering each electrode from the left minus the current exiting to the right. The second $(N-1)$ equations are the energy currents entering each electrode from the left minus exiting to the right. For convergence, these equations should equal zero. A net electrical and heat current is assumed to flow from left to right in the device.

In order to ensure continuity of electrical and heat currents through the device, these $2(N-1)$ equations must be solved simultaneously i.e. the variables should be varied until each equation satisfies the following criteria:

$$|F_i| \leq \text{Tol}, \quad i = 1, 2, 3, \ldots, 2(N-1) \tag{4.7}$$

where Tol is the tolerance of the search.

Because of their exponential components, these equations cannot be solved analytically and so must be solved numerically. The method chosen was the *Newton-Raphson Method* which is iterative and converges quadratically to solve the equations [45]. This method is derived from the Taylor series expansion of a function in one dimension [46]:

$$F(x+\delta) \approx F(x) + F'(x)\delta + \frac{F''(x)}{2}\delta^2 + \cdots \tag{4.8}$$

For small values of $\delta$ (when close to the true solution) the higher order terms become unimportant so for the condition $F(x+\delta) = 0$ (required by Eq. 4.7) we get:

$$\delta = -\frac{F(x)}{F'(x)} \tag{4.9}$$

If we have an initial guess of the variable, $x^i$, an improved value can be given by:

$$x^{i+1} = x^i - \frac{F(x^i)}{F'(x^i)} \tag{4.10}$$

This method is by no means limited to one dimension and can be readily extended to multiple dimensions. For such a system, the updates required for each variable are given by:

$$\boldsymbol{x}^{i+1} = \boldsymbol{x}^i - \frac{\boldsymbol{F}(\boldsymbol{x}^i)}{\boldsymbol{J}(\boldsymbol{x}^i)} \tag{4.11}$$

In order to use this method the *Jacobian Matrix* defined in Eq. 4.12 must be calculated at each iteration.

$$\boldsymbol{J} = \begin{bmatrix} \frac{\partial F_1(\boldsymbol{x})}{\partial x_1} & \frac{\partial F_1(\boldsymbol{x})}{\partial x_2} & \dots & \frac{\partial F_1(\boldsymbol{x})}{\partial x_n} \\ \frac{\partial F_2(\boldsymbol{x})}{\partial x_1} & \frac{\partial F_2(\boldsymbol{x})}{\partial x_2} & \dots & \frac{\partial F_2(\boldsymbol{x})}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_n(\boldsymbol{x})}{\partial x_1} & \frac{\partial F_n(\boldsymbol{x})}{\partial x_2} & \dots & \frac{\partial F_n(\boldsymbol{x})}{\partial x_n} \end{bmatrix} \tag{4.12}$$

The set of equations to calculate the Jacobian were found analytically and generalised and are able to handle a system that has any number of barriers. The partial differential equations used to calculate the Jacobian is shown in Appendix A. These partial differential equations could easily be calculated numerically by using a finite difference approximation such as the Stirling formula [47], but knowing the exact analytic expressions reduces execution time and can increase accuracy.

At each iteration the Jacobian matrix must be inverted in order for the update in Eq. 4.11 to take place. Many of the elements in the Jacobian matrix will be zero because each function to be solved only depends on a small number of variables. By manipulating the order of functions and variables the matrix could be made *band-diagonal*. Special algorithms are available to invert band-diagonal matrices which can decrease execution time and memory usage greatly. Because the number of unknown variables is quite small (18 for a 10-barrier system) this is not necessary as execution time is not prohibitive. In Chapter 6 this will become necessary as a more complex model is derived.

### 4.3.4 Parameters and Boundary Conditions

The barrier height and thermal resistance of each node can be set and are considered parameters of the system. The bias across the last barrier $V_N$ and temperature of the first and last electrodes ($T_0$ and $T_N$) are set. All of the other biases and temperatures of the system are varied using Newton's Method until the continuity equations are solved. Once the system has converged the total bias across the device can be calculated using Eq. 4.17. In fact, once convergence has been reached the real advantages of a numerical system become apparent as any variable of the system can be read out in order to readily achieve such things as temperature and bias profiles across the system.

### 4.3.5 Convergence and Initialisation of Parameters

Although Newton's Method behaves very well and converges quadratically in most cases, it is a well known fact that if the initial conditions of a system are far away from the true solution then the system may not converge at all or may converge to a local, rather than global, solution. This is equivalent to saying that $\delta$ must be small in Eq. 4.8.

For given boundary conditions, $T_0$, $T_N$ and $V_N$, the system variables are initialised in the following manner:

$$V_i = V_N \qquad i = 1, 2, 3, \ldots, N-1 \tag{4.13}$$

(the bias across each barrier is set equal to the bias across the last barrier) and

$$T_i = T_c + \frac{(T_h - T_c)i}{N} \qquad i = 1, 2, 3, \ldots, N-1 \tag{4.14}$$

(the temperature of each electrode is linearly spaced between the two end temperatures). These initial conditions generally ensure convergence.

## 4.4 Comparison Between Single- and Multiple-Barrier Devices

### 4.4.1 Heat Current

One very interesting outcome of using multiple barrier devices rather than single barrier devices is that the heat current leaving the cold side no longer saturates at large applied

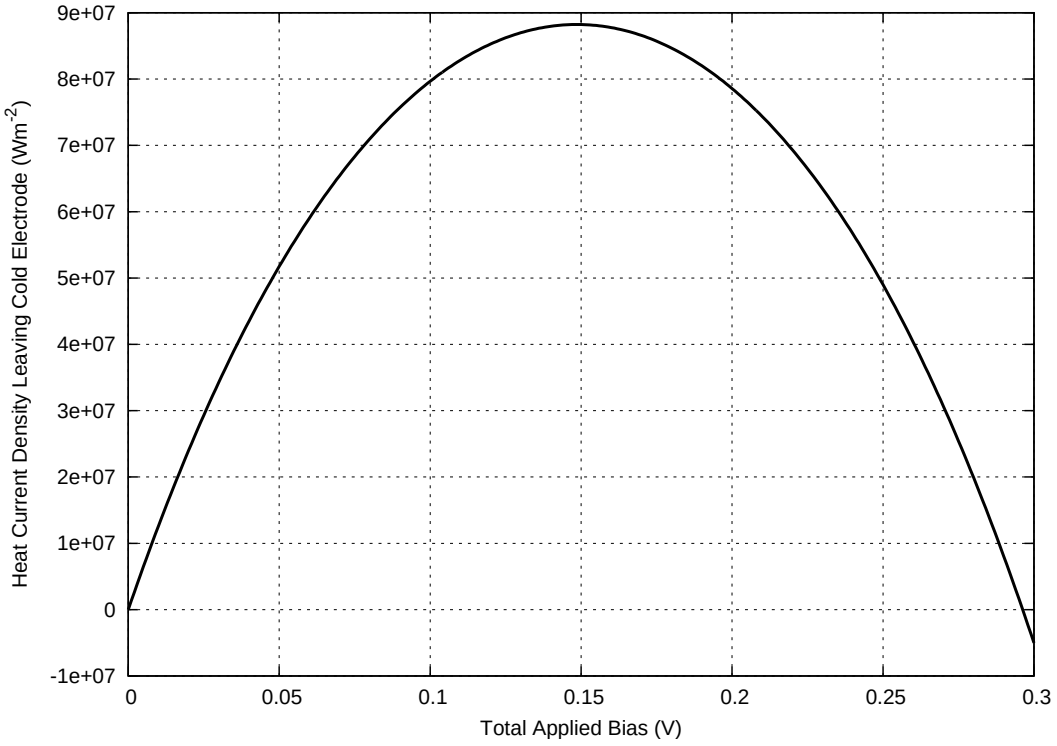


Figure 4.3: Heat current density leaving cold electrode for 10-barrier device:
$\phi_i = 0.1$ V, $T_c = 300$ K, $T_h = 300$ K, $R_i^{th} = 41.14 \times 10^{-9}$ m$^2$·K/W, $m^* = 1, N = 10$.

bias as in Fig. 3.3. On the contrary, the net heat current removed from the cold side actually begins to decrease as the bias increases past a critical voltage. Because of the complexity of the system this critical voltage cannot be calculated analytically and can only be found numerically for a given system.

Fig. 4.3 shows the heat current leaving the cold electrode for a $10$-barrier device. The device parameters are as follows:

$\phi_i = 0.1$ V, $T_c = 300$ K, $T_h = 300$ K, $R_i^{th} = 41.14 \times 10^{-9}$ m$^2$·K/W, $m^* = 1, N = 10$.

The total bias across the device, $V_{tot}$, is varied from 0 V to 0.3 V in 0.001 V increments. The parameters used here are somewhat arbitrary but are chosen to highlight trends in device performance. The same behaviour will be apparent with other choices but may not be as obvious. Real material parameters will be used and compared in section 4.6.2.

As stated in section 4.3.4, only the bias across the last barrier can be set and all other biases in the system are varied to solve the system equations. When interested in the total

bias across the device (as in this case), Newton's Method is once again employed (this time in 1-dimension) to find the end bias, $V_N$, that results in the desired total bias, $V_{tot}$, across the device. The equation that is solved is:

$$V_{tot} - V_{Actual}(V_N) \leq \text{ Tol} \tag{4.15}$$

where $V_{Actual}$ is a function of $V_N$.

Referring to Fig. 4.3 it is found that the maximum heat current removed from the cold side occurs at an applied bias of around 0.15 V for this device. After this the heat current removed begins to decrease and actually becomes negative near 0.3 V. In other words, at a large applied bias heat actually starts to be deposited into the cold electrode from the second electrode.

This can be explained by referring to Fig. 4.4 which shows the same device at an applied bias of 0.3 V. A profile of the bias of, temperature of, and heat current in each electrode is plotted. Also shown is the work done across each barrier. The bias profile shows that a positive bias is applied across each electrode, which ensures Mode 1 operation — Class 1 equations are valid. The temperature profile shows how the counterintuitive result of heat current entering the cold electrode can occur.

In the single-barrier case both electrodes are considered connected to an infinite thermal reservoir and so their temperatures are held constant. In the multiple-barrier case only the end temperatures are set and held constant. As the bias across the device increases so too does the net electrical current in the device. Work equal to $J^{net}V_1$ is deposited into the second electrode in addition to the heat current leaving the first electrode. This continues throughout the device with the net heat current entering each electrode being larger than the previous electrode due to the work done in the barrier, as shown in the last plot of Fig. 4.4. So even though a small heat current enters the cold electrode from the right, a much larger heat current leaves the hot electrode due to the work done in the system. It is seen that the work done in each barrier is equal to the increase in heat current from one electrode to the next.

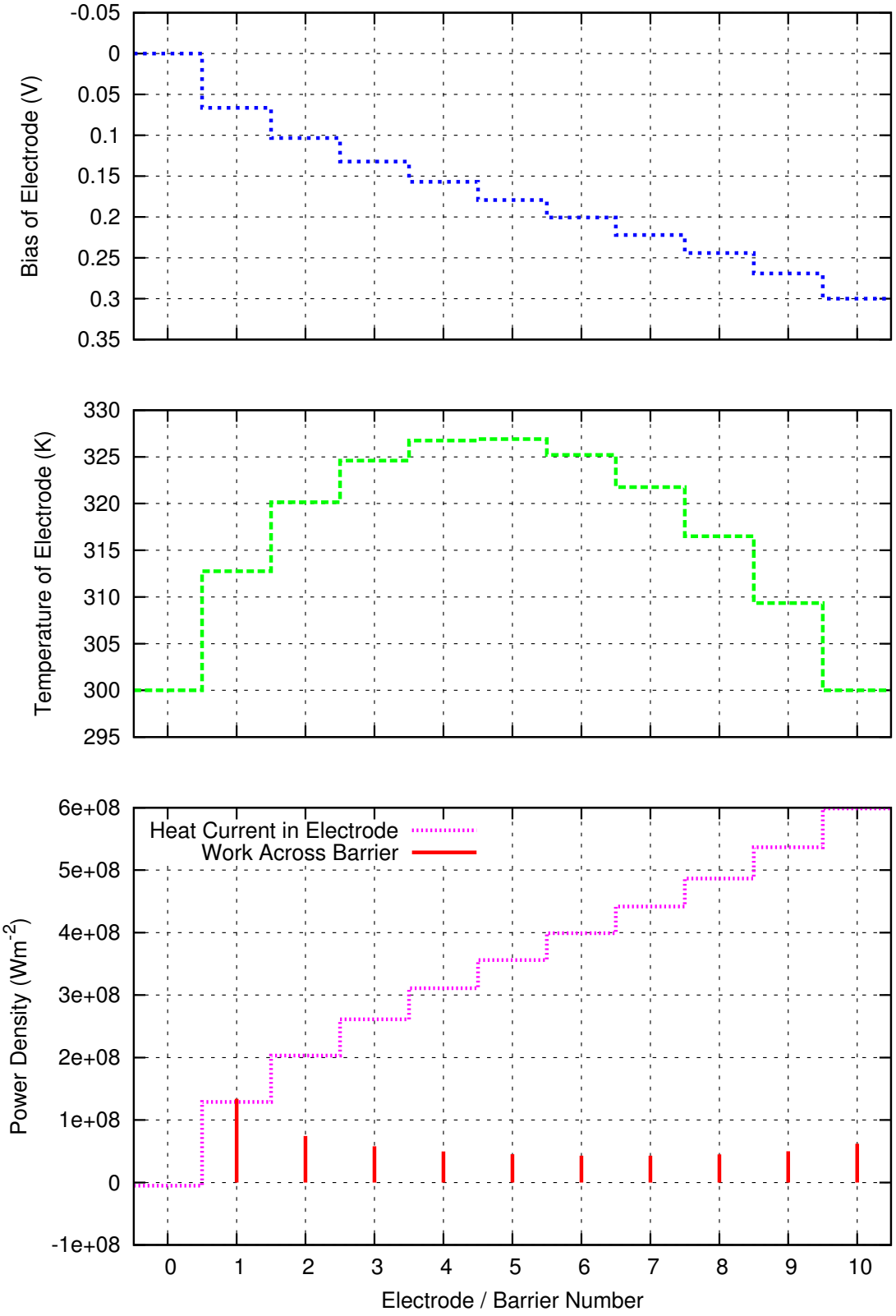


Figure 4.4: Profile of 10-barrier device, Total Applied Bias $V_{tot} = 0.3$ V:
$\phi_i = 0.1$ V, $T_c = 300$ K, $T_h = 300$ K, $R_i^{th} = 41.14 \times 10^{-9}$ m$^2$·K/W, $m^* = 1$, $N = 10$.

Because each inner electrode is considered to be thermally isolated except for the thermionic and conduction heat transport from one electrode to the next, the inner electrode temperatures are increased in order to satisfy the continuity equations.

At first inspection this does appear to be a waste of power as all this work is done on the system whilst achieving no cooling and causing a large temperature gradient between the first and second electrodes. The example shown here is an extreme case and would not be considered in a real device. It will be shown that under correct operating conditions a multiple-barrier device will still perform better than a similar single-barrier device, even when there is an increases of intermediate electrode temperatures.

### 4.5 Efficiency of Multi-barrier devices

Once the equations have been solved for a given set of parameters (such as work function and barrier thickness), the efficiency of each device is calculated in order to compare performance of different systems. The efficiency of the device is the heat taken from the cold side divided by the work needed to do so and is given by:

$$\eta = \frac{J_{qc}}{J_c V} \tag{4.16}$$

where

$$V = \sum_{i=1}^{N} V_i \tag{4.17}$$

The motivation for moving from a single- to multi-barrier system was to overcome decreasing device efficiency due to heat conduction and to ensure that the devices are operating thermionically (i.e. electrons move ballistically across each barrier). It was proposed that having a number of barriers would decrease $\Delta T$ across each barrier and so conduction heat conduction (the last term in Eq. 4.3) would be reduced at each layer. Simulation results have shown that while $\Delta T$ is reduced at each barrier, the overall efficiency of the device does not increase as the number of barriers increases. By slightly modifying the model used (making $R$ dependent on $N$) it is shown that the multi-barrier system is more efficient than a single-barrier system.

As an alternative to decreasing $\Delta T$ across each barrier, the conduction heat current can be decreased by increasing the effective thermal resistivity, $R$, across each barrier. The multi-layer cooler may be more efficient for different reasons than proposed. Chen [43] has shown that the thermal conductivity of a superlattice decreases as the number of interfaces increases.

### 4.6 Maximum Cooling Power

In modelling these devices it is important to find which combination of parameters will give the greatest cooling power. When the hot side is held at a constant temperature (it is thermally connected to an infinite thermal reservoir) this corresponds to finding the combination of parameters that bring the cold side to a minimum temperature.

So far the two end temperatures of the devices have been held constant in order to inspect device operation between these two temperatures — we have been interested in how the devices perform when operating between two set temperatures.

If we hold the hot temperature and allow the cold side temperature to vary we can find the minimum cold temperature possible for a given device. To determine this we start with $T_c = T_h$. In the general case both $T_c$ and $T_h$ will vary when a bias is applied across the device because of heat currents leaving and entering the two end electrodes. For the case when the hot side temperature is considered constant, $T_c^{min}$ is found simply by determining when the heat current leaving the cold side is zero. If a heat current leaves the cold side, then obviously the cold side has not reached its minimum temperature. On the other hand, the only way a heat current can enter the cold side from the second electrode is if it has been artificially held at lower temperature. This reasoning is only valid if it is assumed that the cold temperature is thermally isolated from its surroundings. In a real device as soon as the cold electrode is connected to a hotter surface (i.e. the surface that requires cooling), a net heat current will flow into the cold electrode. For the purpose of finding the minimum cold electrode temperature achievable it is considered that the net

heat current entering the cold electrode is always equal to the net heat current leaving this electrode.

Considering just the first two electrodes of a multi-barrier device, when a bias is initially applied heat current due to thermionic emission will leave the first and enter the second electrode. This will cause the first electrode to cool and the second to heat. This will increase the temperature gradient across the first barrier and thus the backward heat flow due to conduction will also increase. This will, in turn, increase the temperature of the first electrode and decrease the temperature of second electrode. This reduces both the net heat current due to thermionic emission and the heat current due to conduction. This process will continue until the net heat current leaving the cold side is zero. Thus the minimum cold temperature for that bias will be achieved.

### 4.6.1 Minimum Cold Electrode Temperature vs. Applied Bias

The cold temperature of the device can be varied at a set applied bias to find when the net heat current leaving the cold electrode is zero. This is achieved by using a Newton-Raphson scheme in 1-dimension with $T_c$ now the variable and $V_N$ held constant. By doing this for a range of biases the cold electrode temperature dependence on applied bias can be found.

Fig. 4.5 shows variation of cold electrode minimum temperature with applied bias for the same device parameters as used in Fig. 4.3. Both a single- and 10-barrier device results are shown. The multiple-barrier device can achieve a much lower cold electrode temperature.

As is expected a negative applied bias increases the heat flow from hot- to cold- electrode and thus causes an increase in the cold electrode temperature. A positive applied bias initially causes the cold electrode of the devices to decrease as heat is removed. For the single barrier device this temperature plateaus as the heat current removed from the cold side saturates. For the 10-barrier device a global minimum cold electrode temperature is

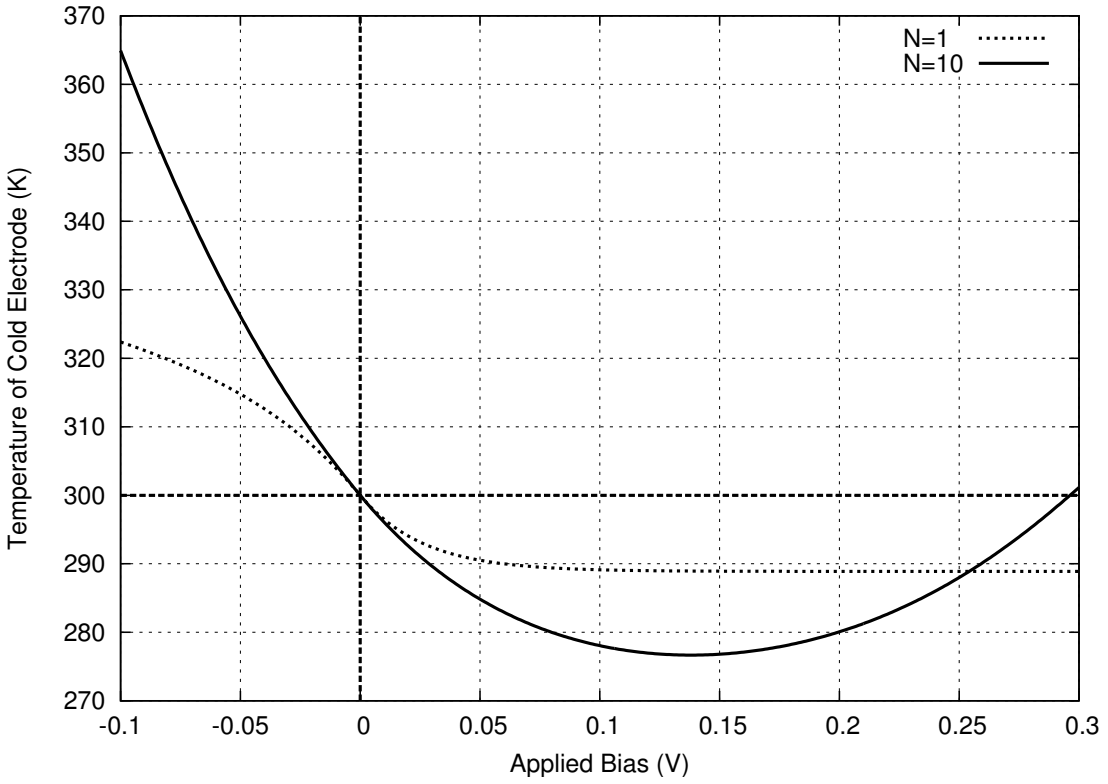


Figure 4.5: Minimum Cold Electrode Temperature vs. Applied Bias :
$\phi_i = 0.1$ V, $T_h = 300$ K, $R_i^{th} = 41.14 \times 10^{-9}$ m$^2$·K/W, $m^* = 1, N = 1, 10$.

achieved at an applied bias of around 0.14 V. As the applied bias is further increased the temperature begins to increase again due to the heat current decreasing, as discussed in section 4.4.1.

### 4.6.2 Device Parameters that induce the greatest temperature gradient

By using a 1-dimensional Newton-Raphson scheme, the bias across the device can be varied until the minimum cold temperature is achieved for a given set of device parameters.

Table 4.1 gives some selected device parameters of common semiconductors. These values are used to calculate two important operating parameters: the mean-free path and maximum thermal resistance per unit area (which is dependent on the mean-free path). These parameters in are presented in Table 4.2. In the tables, (n) and (p) refer to electrons

Table 4.1: Selected Material Parameters

| **Material** | **m* ($m_e$)** | **$\mu$ (cm$^2$/V·s)** | **Therm. Resistivity (K·cm/W)** |
|---|---|---|---|
| GaAs (n) [42] | 0.063 | $8.5\times10^3$ | 1.82 |
| GaAs (p) [42] | 0.51 | $4.0\times10^2$ | 1.82 |
| $Al_{0.07}Ga_{0.93}As$ (n) [42] | 0.069 | $7.2\times10^3$ | 2.42 |
| InSb (n) [40] | 0.014 | $7.7\times10^4$ | 5.56 |
| InSb (p) [40] | 0.43 | $8.5\times10^2$ | 5.56 |
| InP (n) [42] | 0.08 | $5.4\times10^3$ | 1.47 |
| InAs (p) [42] | 0.023 | $4.0\times10^4$ | 3.70 |
| $Ga_{0.47}In_{0.53}As$ (n) [42] | 0.041 | $1.2\times10^4$ | 20.0 |

Table 4.2: Calculated Material Parameters

| **Material** | **Mean Free Path (nm)** | **Max. R$^{th}$ ($\times$10$^{-9}$ K·m$^2$/W)** |
|---|---|---|
| GaAs (n) | 81.8 | 1.49 |
| GaAs (p) | 11.0 | 0.20 |
| $Al_{0.07}Ga_{0.93}As$ (n) | 72.0 | 1.74 |
| InSb (n) | 349.3 | 19.42 |
| InSb (p) | 21.4 | 1.19 |
| InP (n) | 58.6 | 0.861 |
| InAs (p) | 232.6 | 8.615 |
| $Ga_{0.47}In_{0.53}As$ (n) | 93.2 | 18.63 |

and holes being majority carriers, respectively. This can be achieved by either doping the material with donors or acceptors, or by surrounding the barrier material with highly doped regions.

n-type InSb and n-type $Ga_{0.47}In_{0.53}As$ have much higher maximum thermal resistance per unit area than the other materials. n-type InSb, though, has a comparatively low effective mass. A reduced thermal mass means that, on average, less energy is transported per electron contributing to thermionic current. n-type $Ga_{0.47}In_{0.53}As$ has a larger effective mass although it is not as large as for materials where holes are the majority carriers. These two materials will be of interest for use in devices. It should be noted that $Ga_{0.47}In_{0.53}As$ has a much larger thermal resistivity than that of the related binaries (InAs and GaAs). This can be attributed to lattice disorder due to the random distribution of constituent atoms in

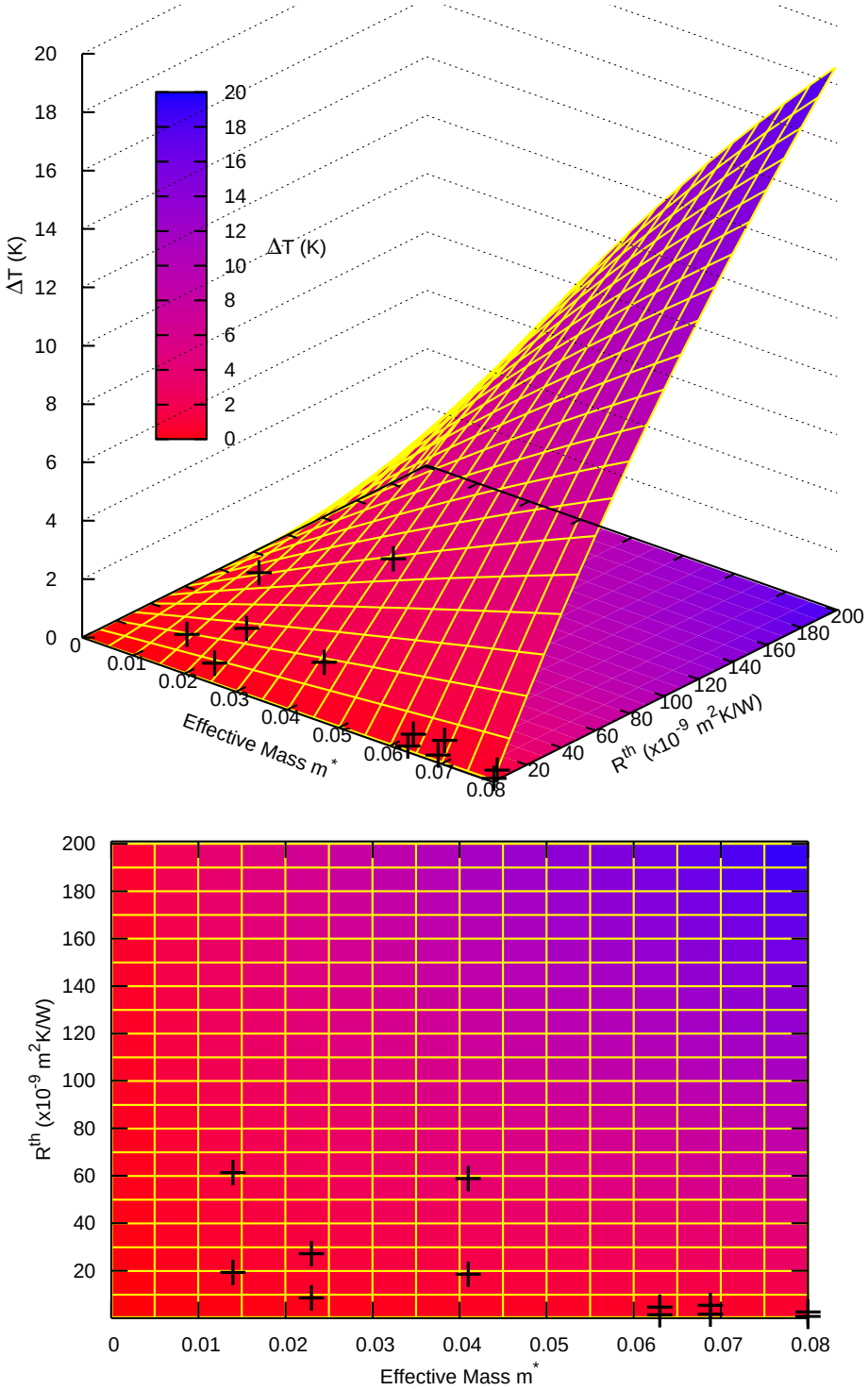


Figure 4.6: Maximum Cooling of Cold Electrode vs. Device Parameters. $\phi_i = 77$ meV, $T_h = 300$ K, $N = 10$.

the two sublattice sites [48].

With its relatively large effective mass, mobility and thermal resistivity, n-type $Ga_{0.47}In_{0.53}As$ will be the material of most interest. Comparing with n-type InSb, it has a much larger thermal resistivity, but due to is increased effective mass and reduced mobility (and therefore reduced mean-free path), the maximum thermal resistance per unit area for $Ga_{0.47}In_{0.53}As$ is slightly less than InSb, although still much larger than the other materials. Fig. 4.6 is used to investigate the interplay between effective mass and thermal resistance of materials and how they effect maximum cooling.

Fig. 4.6 shows the maximum cooling for 10-barrier structures operating with the hot electrode at 300 K. By trying different barrier heights it was found that the cooling power increases substantially with decreasing barrier height (the same is not necessarily true for efficiency). A barrier height of 77 meV is used for the devices in Fig. 4.6. This value is the minimum value for which Maxwell-Boltzmann statistics are still valid in the model (see section 2.3.1). As discussed in section 3.5.1 the effective mass and thermal resistance of materials are two important parameters that effect device performance. Due to the effective mass appearing in the expression for mobility, and therefore mean-free path and maximum thermal resistance per unit area, a numerical solution is used to compare different materials.

By trying different numbers of barriers it was found that there is substantial increase in device performance between 1- and 10- barrier systems. Increasing the number of barriers further, to say 20 or 50, does not substantially increase the performance of the devices but does add to the complexity of their fabrication.

The values calculated in Table 4.2 are shown in Fig. 4.6 to compare the performance of ‘real’ devices. Some values are not shown (particularly those for p-type materials) because, due to conduction heat flow (as discussed in section 3.5.2), they can never operate as a refrigerator at room temperature. p-type materials generally have a higher carrier

effective mass but this is offset by reduced mobility.

### 4.6.3 Discussion of Results

Table 4.3: Maximum cooling for different materials - $^*$: increased thermal resistivity due to phonon scattering

| **Material** | **Max.** $\Delta T$ **(K)** | **Max.** $\Delta T^*$ **(K)** |
|---|---|---|
| GaAs (n) | 0.135 | 0.426 |
| $Al_{0.07}Ga_{0.93}As$ (n) | 0.182 | 0.573 |
| InSb (n) | 0.413 | 1.289 |
| InP (n) | 0.105 | 0.331 |
| $Ga_{0.47}In_{0.53}As$ (n) | 1.147 | 3.511 |

Table 4.3 summarises the results for selected materials shown in Fig. 4.6. The second column shows the maximum cooling possible for a 10-barrier structure made of the material in column one, operating at a temperature of 300 K with a barrier height of 77 meV. The third column shows the maximum cooling for the same material, but with the thermal resistivity of the material increased due to phonon scattering. This mechanism is approximated by multiplying the bulk resistivity by the square root of the number of barriers. This is done in accordance with the increase in thermal resistivity of superlattices modelled by Chen [43] and as discussed in section 4.1.2.

The devices used in this work are GaAs/AlGaAs based systems. The Al concentration is around 7% in the heterostructure barriers. The maximum cooling for a 10-barrier system is expected to be around 0.182 K if bulk thermal resistivity is used and 0.573 K if an increase in phonon scattering is assumed. Because the barriers are made from GaAs the maximum cooling is expected to lie between the value for the two different values i.e. between 0.135 and 0.182 K if bulk resistivity is used or 0.426 and 0.573 K if increased thermal resistivity is assumed.

## 4.7 Conclusion

This chapter has seen the extension of the single-barrier device equations of chapter 3 into a system that can handle any number of barriers. The advantages of using a multiple barrier system were confirmed by the fact that considerably more cooling power is achievable whilst only having a slight loss in device efficiency. By having only around $\Delta T/N$ temperature drop across each barrier (as opposed to $\Delta T$ for a single-barrier system) more heat is removed from the cold side of the device due to reduced heat conduction across each barrier. In addition to this, further heat is removed when phonon scattering due to interfaces between materials is taken into account. This results in greater cooling than for single-barrier devices.

Due to the work done in transporting carriers across each barrier it was found that the heat transport increases throughout the device. In some cases this increases the temperature of intermediate electrodes above the temperature of both the end electrodes. This appears to be a waste of energy but multi-barrier systems still out perform single barrier systems. If this heat could be removed by some other means (such as optical emission) the performance of the devices would be increased further.

By employing a numerical solution to the problem many advantages of such a system become apparent. Once convergence of the system equations has been reached, the performance of devices can be modelled under a multitude of different situations. For instance, the maximum cooling for a given applied bias can be found by varying the temperature of the cold side until the heat current leaving the first electrode is zero. Likewise, the maximum cooling for a combination of device parameters can be found by repeating this step for a range of applied biases until the minimum cold temperature is achieved.

Due to increased average energy of carriers it is found that these devices operate more efficiently at elevated temperatures. This means that these devices may have applications at high temperature, such as running the devices in reverse as power generators.

Due to the effective mass of carriers appearing in the numerator of Richardson's constant, increased effective mass generally means increased cooling power and efficiency. But the effective mass also appears in the equation for mean-free path of carriers which in turn sets the maximum thermal resistance per unit area possible for different materials. Because of the complex interplay between effective mass and maximum thermal resistance, graphical methods were used to compare the theoretical minimum temperature for 10-barrier devices. For the devices used in this work, cooling of less than half a degree is expected.

The next chapter will present the design of devices and experiment setup and results for the devices used in this work.

# CHAPTER 5

# DEVICE DESIGN AND EXPERIMENTAL SETUP

## 5.1 Introduction

This chapter discusses device design, fabrication and experimental techniques used to characterise the devices. All devices in this chapter are 10-barrier devices made using GaAs/AlGaAs heterostructures.

## 5.2 Motivation for Using GaAs/AlGaAs Heterostructures

GaAs is by no means an ideal material to use for these thermionic devices, as seen by the previous chapter. GaInAs would be a better material, but due to availability GaAs/AlGaAs was used. Our first-generation of devices were made by the Australian National University before detailed modelling had been completed and this material was readily available at the time.

## 5.3 Device Design

### 5.3.1 First- and Second- Generation Device Design

The first-generation devices where produced when the equations were not entirely understood by the author. In these devices barrier heights were considered to be the conduction band offset in the devices. On closer inspection of the derivation of Richardson's equation it is clear that barrier heights are not measured from conduction band discontinuities, but rather are measured from the chemical potential in the electrode to the conduction band in the barrier. On modelling the $I$-$V$ characteristics of experimental data for these devices it was found that the actual barrier heights were much larger than expected. The conduction band offset, rather than the barrier height measured from the chemical potential, was set

to a nominal 77 meV.

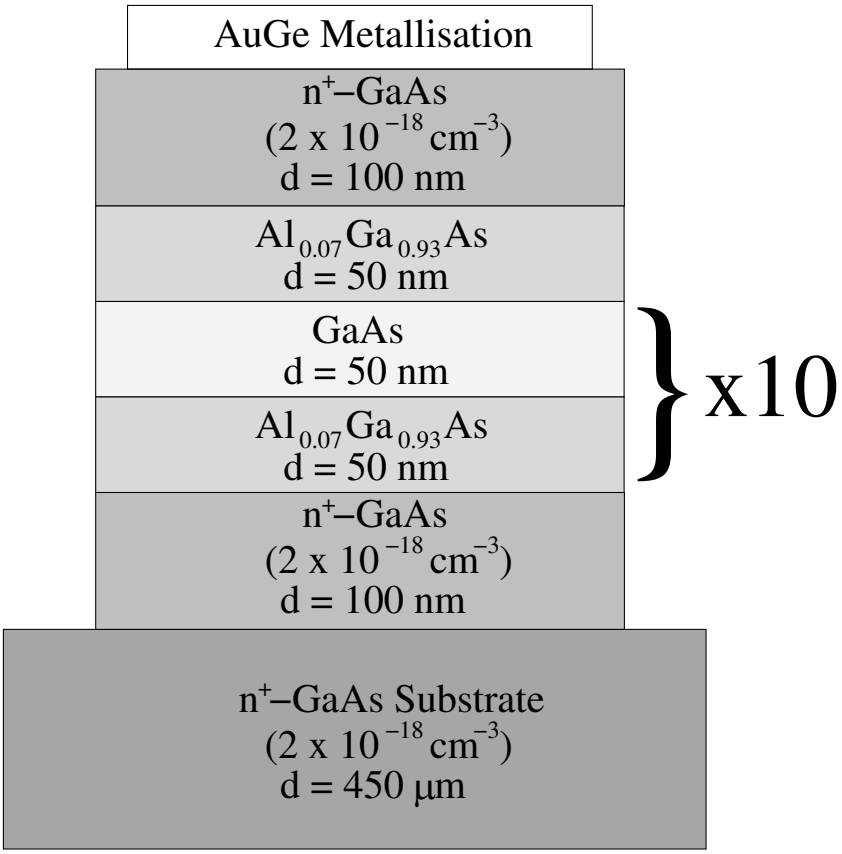


Figure 5.1: Structure of First-Generation Device

First- and second- generation devices have exactly the same design. The only difference between them is that the second-generation sample also had metal contacts put directly onto the substrate. This was done so that the effect of the substrate and top- and bottom-contacts could be measured separately. Initially it was not know exactly how these would effect the experimental results (energy dissipation, impedance) and so it was expected that by measuring the characteristics of the substrate/metallisation alone would allow isolation of the devices.

Chapter 6 sees the development of a better model for the devices and it will be seen that the substrate and metallisation have little effect on the overall experimental results.

### 5.3.2 Third-Generation Device Design

With the experimental results showing that the barrier height for the first- and second-generation devices was too large, third-generation devices were designed to address this problem.

The first two generations of device consisted of undoped GaAs/AlGaAs heterostructures. Because of this the barrier height was much larger than the nominal 77 meV. By highly doping the active region of the third-generation devices the chemical potential of the devices was moved closer to the conduction band of the well material (GaAs). By doing so the barrier height can be engineered to be closer to the required value.

The simulation used to design the third-generation devices has not yet been discussed. Chapter 6 introduces analysis of the $I$-$V$ characteristics of the first two generations of the device. It will be shown that the model developed so far does not sufficiently describe these experimental results. A new model is developed which does model the results more successfully and will be used to analyse the first two generations in Chapter 7. This chapter will then use the model to design the third-generation devices and further discuss the motivation for the new device design. This section is included here to introduce the third-generation devices as the experimental results will be presented in this chapter, although the actual device design will not be discussed until Chapter 7.

## 5.4 Experimental Setup

### 5.4.1 Electrical Measurements

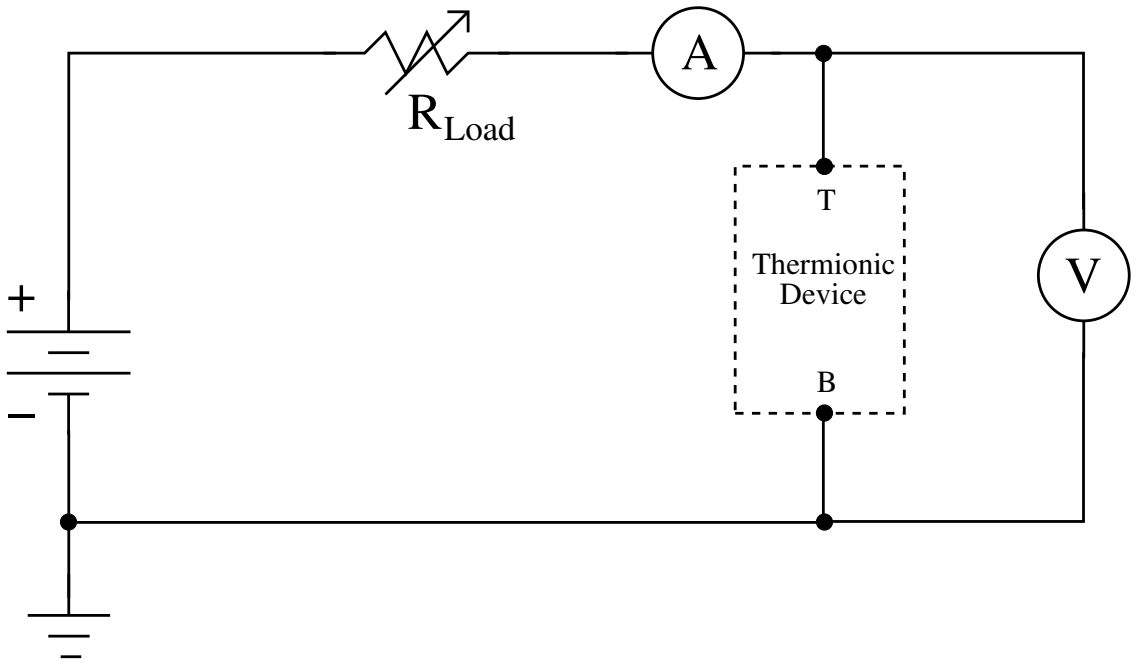


Figure 5.2: Electrical Schematic of Experimental Setup showing contacts to (T)op and (B)ottom of Device

Fig. 5.2 shows the electrical schematic of the simple circuit used to measure the $I$-$V$ characteristics of the device and to also input work into the system for thermal measurements. The electrical contacts to the (T)op and (B)ottom of the device are shown in the experimental setup in Fig. 5.3. Referring to this figure the samples are placed on a brass heat sink. No metal contacts were put onto the bottom of the samples during fabrication. InGa eutectic was used to promote electrical contact between the devices and the brass heat sink.

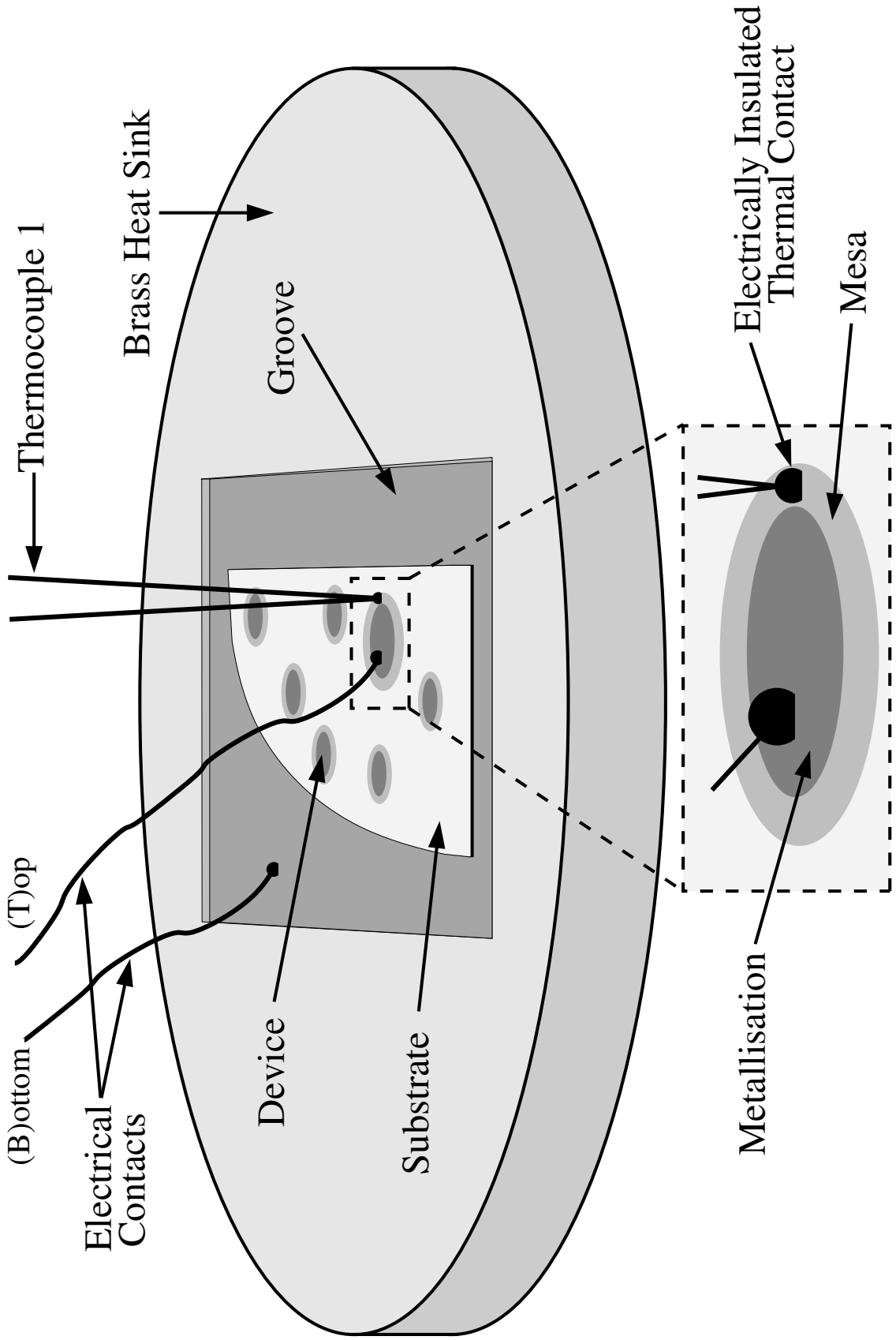


Figure 5.3: Experimental Setup for measuring $I$-$V$ and contact thermal characteristics

## 5.5 Experimental Results

### 5.5.1 First- and Second- Generation $I$-$V$ Characteristics

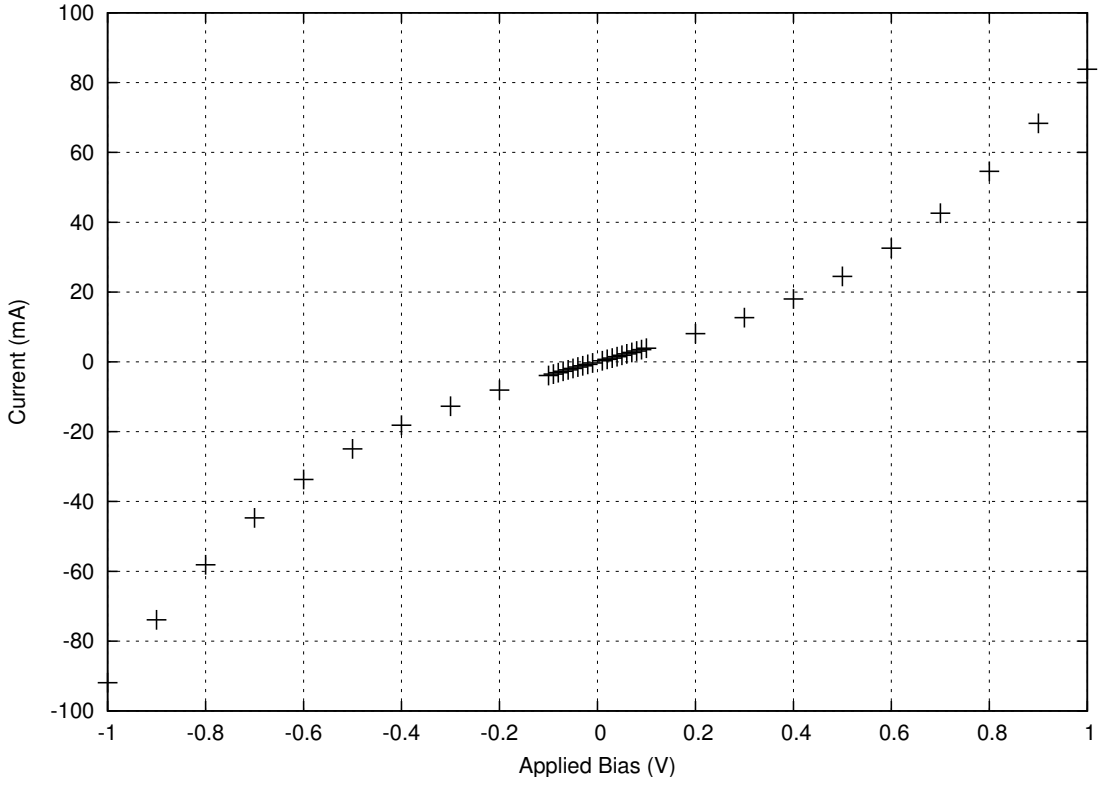


Figure 5.4: Typical I-V characteristics of 6 mm first- and second- generation device

The majority of experimental work on the first- and second- generation devices was carried out by Sueping Lee and is presented extensively in the Master's Thesis 'Multilayer thermionic cooling in GaAs-$Al_xGa_{1-x}As$ heterostructures' [49]. Some selected results will be taken and analysed in Chapters 6 and 7 to determine the accuracy of the models used. Fig. 5.4 shows typical results for a first- or second- generation device with diameter 6 mm.

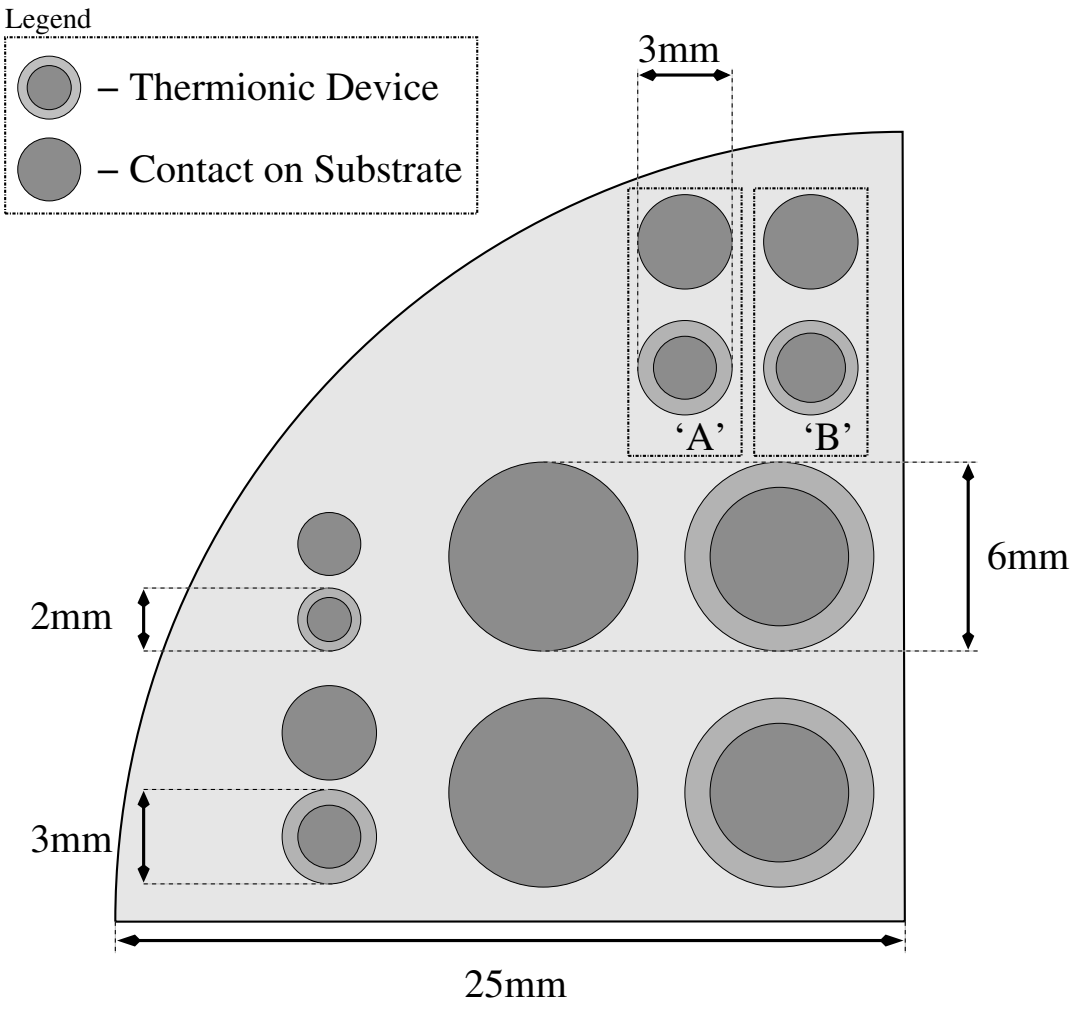


Figure 5.5: Third-generation devices and metal contacts on $n^+$-GaAs substrate

### 5.5.2 Third-Generation $I$-$V$ Characteristics

Fig. 5.5 shows the layout of the third-generation sample used. Electrical characteristics for the two 3 mm devices ‘A’ and ‘B’ will be presented and compared to show repeatability of the results. As shown on the figure, the third-generation sample has both devices and metal contacts only on the substrate. This allows determination of $I$-$V$ characteristics of the devices and metallisation alone.

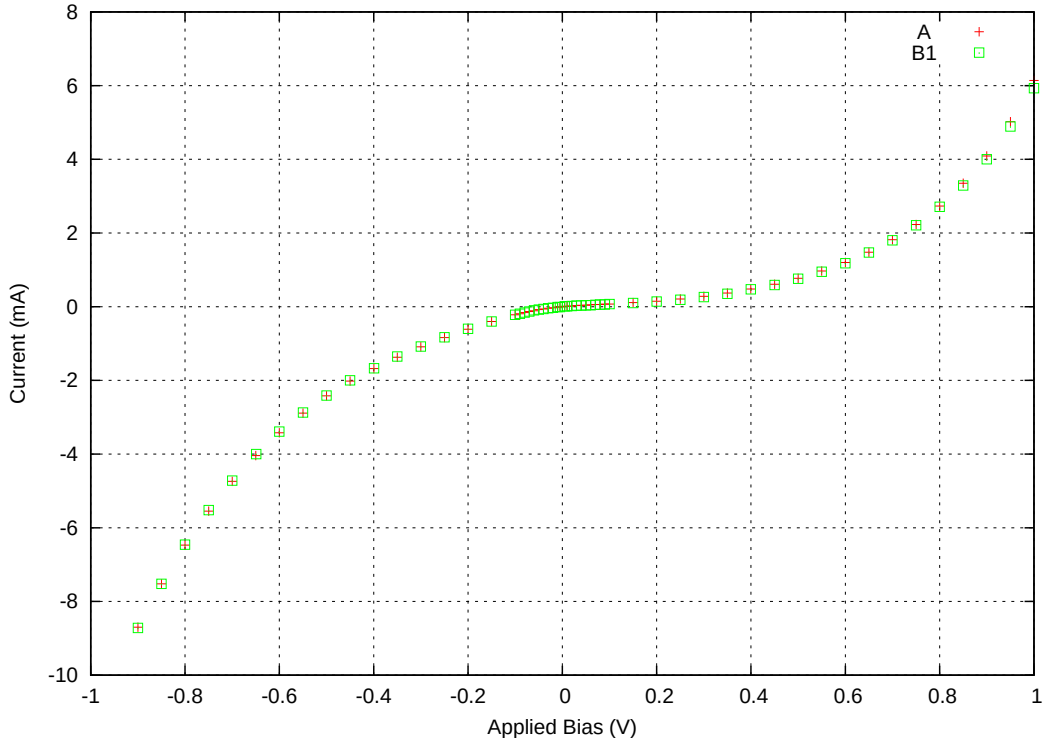


Figure 5.6: $I$-$V$ Characteristics of two third-generation devices

Fig. 5.6 shows $I$-$V$ characteristics of two different samples, ‘A’ and ‘B’, taken at the same time at the same temperature. As can be seen there is very good agreement between the two samples.

Fig. 5.7 shows $I$-$V$ characteristics of sample ‘B’ taken on two separate occasions at slightly different temperature (equal to the ambient temperature of the room). The results taken at the slightly higher temperature have a higher current due to more carriers being available to contribute to drift-diffusion current and to overcome the potential barriers in

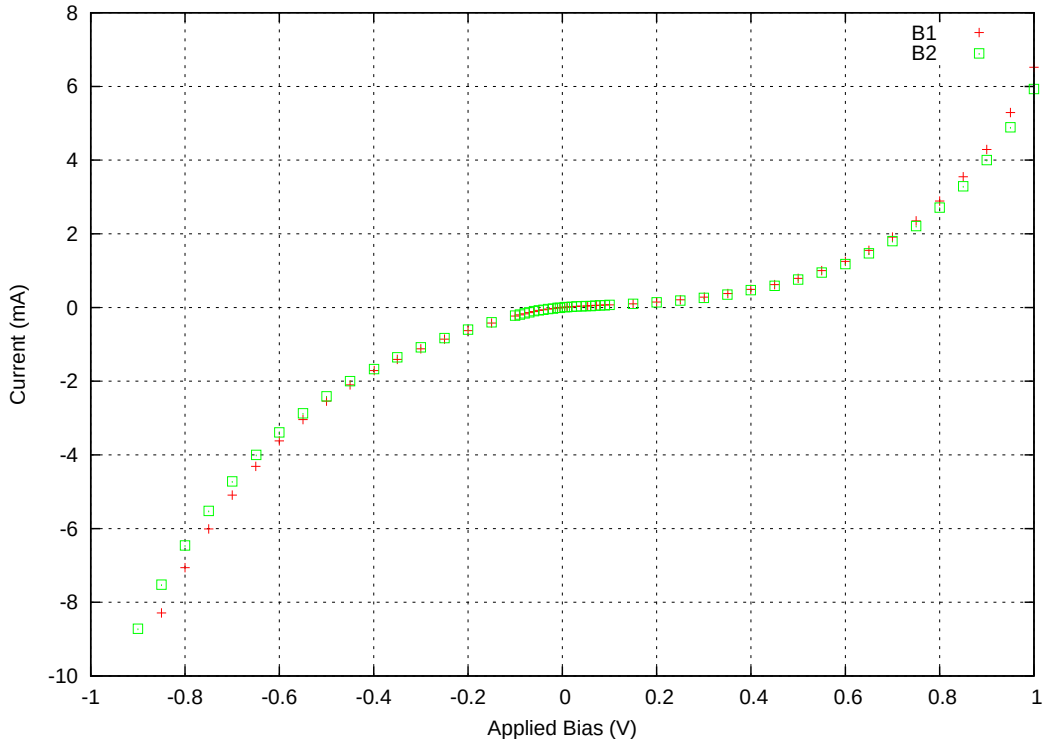


Figure 5.7: $I$-$V$ Characteristics of a third-generation device at different temperatures

the device. For all results the current is larger for a negative applied bias than for positive applied bias.

Fig. 5.8 shows the $I$-$V$ characteristics of the metal contacts directly on the substrate. Once again all results are similar and the current is larger for negative applied biases.

The experimental result presented here will be analysed in chapters 6 and 7.

## 5.6 Non-Contact Temperature Determination

### 5.6.1 Introduction

Because of the thermal mass of the thermocouples used for temperature measurement, their physical contact with the samples may influence thermal measurements. If any cooling occurs in the samples it may only be obvious for a short time before joule heating in the substrate, contacts or device itself causes the temperature of the whole sample to increase, even when thermally connected to the brass heat sink. To alleviate this problem

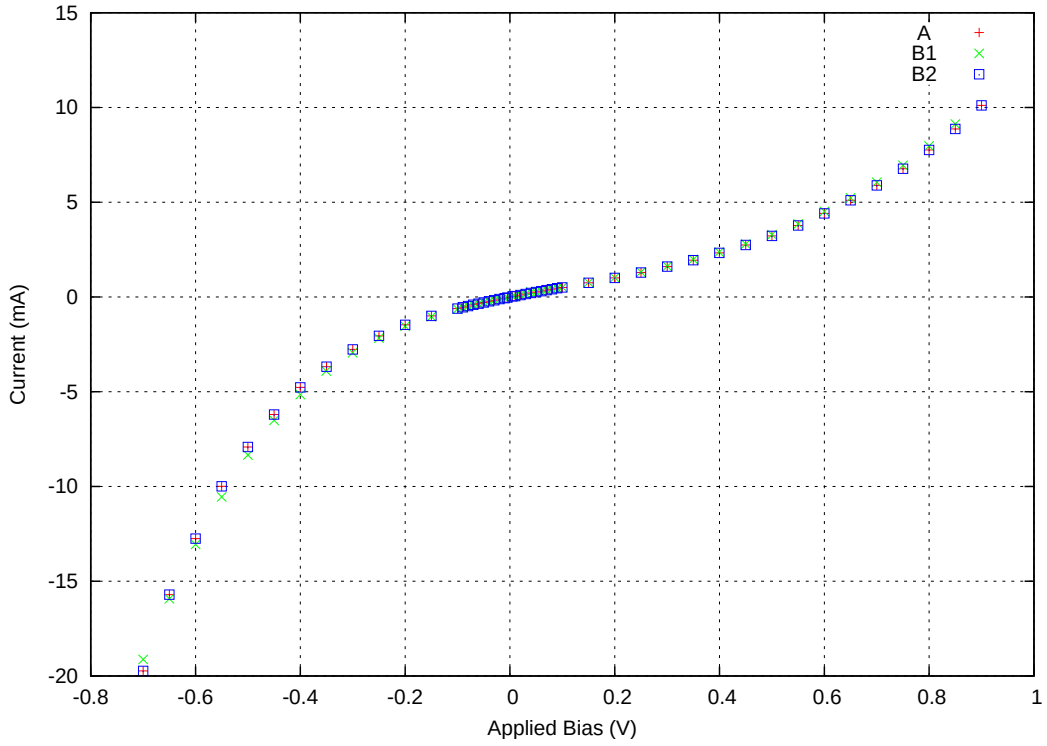


Figure 5.8: $I$-$V$ Characteristics of a third-generation metal contact on substrate

work has begun on using non-contact optical methods for temperature measurement.

Reflectance measurements rather than transmission measurements are needed because transmission measurements will measure temperatures within the device rather than on the surface of the device as required. Due to joule heating in the device and energy transfer away from the surface the temperature inside the device is expected to be higher than on the surface. Any cooling occurring at the surface will not be noticeable due to this internal heating.

### 5.6.2 Optical Modulation Spectroscopy

Several non-contact methods are available for analysing devices in a non-destructive way. Some examples are electroreflectance, magnetoreflectance, stress reflectance and photoreflectance. Modulated photoreflectance is a common technique for determining band-gap energies, doping concentrations and alloy compositions in samples [50]. Most of this

work is carried out a low temperature.

One technique that has shown promise at elevated temperatures (i.e. near room temperature) is differential reflectance. This method relies on the spatial inhomogeneity of samples as the source of modulation. By not relying on an external source of modulation to generate the optical signal this method has advantages because external modulation sources (such as strain) can become unreliable at elevated temperatures [51].

Differential reflectance involves periodically measuring the reflectivity difference between two neighbouring points on the sample as a function of photon energy. As stated this method relies on imhomogoneity of the sample surface to generate the measured signal. This imhomogoneity can be introduced by applying a temperature gradient or electric field across the sample [52]. By determining how the optical signal changes with temperature it is hoped that this method may provide a reliable method for non-contact temperature measurement.

Very preliminary work was carried out in this thesis to study the feasibility of using optical non-contact measurements to determine temperature changes in the devices. The work in effect measured transmission spectra in spite of a “reflectance” geometry and as such is still far from providing the technique outlined above. It was carried out more to determine the capabilities of the digital spectrometer used. Further work is to be carried out in developing the differential reflectance technique needed.

### 5.6.3 Experimental Setup

To carry out this experimental work a Stellarnet EPP2000 Multichannel Digital Spectrometer was used. This spectrometer measures a range of frequencies in the optical spectrum simultaneously with a resolution of 0.4 nm. It is interfaced to a computer for data collection and analysis.

A very simple setup was used for these measurements. A diffuse ‘white’ light source such as an incandescent or halogen lamp was shone onto the samples at an angle equal to

the angle of the spectrometer probe.

For these measurements bulk GaAs and InP samples were used. The samples were placed onto a heater and measurements were made over a range of temperatures to determine if any noticeable changes in the spectra were apparent. Optical noise due to reflection off surrounding surfaces was reduced by placing the samples on a dark surface.

### 5.6.4 Experimental Results

The experimental results for bulk GaAs and InP are shown in Fig. 5.10. Analysis of the results show that the spectra result from transmission rather than reflectance of incident light. Although these results will be of little use in temperature determination of the real devices they do illustrate the potential use of this digital spectrometer.

The temperature range was swept from 303 K (room temperature) to 363 K in 5 K increments. The transmission spectrum illustrates how the band-gap of the materials changes with temperature. Using de Broglie relations [33]:

$$\lambda_i = \frac{hc}{E_{bg}} \tag{5.1}$$

to relate the band-gap energy, $E_{bg}$, of the material to the wavelength of the incident light, $\lambda_i$, it is found that the band-gap energy of the two materials at 300 K are equivalent to the following wavelengths:

$$\begin{aligned} E_{bg}^{GaAs} &= 1.424 \text{ eV} \equiv 871 \text{ nm} \\ E_{bg}^{InP} &= 1.344 \text{ eV} \equiv 923 \text{ nm} \end{aligned}$$

Any incident light with shorter wavelengths than the band-gap is absorbed whilst light with a wavelength longer than this is transmitted.

The maximum count of incident light around the band-gap of GaAs changes by around 100 when the temperature changes by 60 K, meaning there is an average change of around 1.7 counts per degree. For InP this value is around 1.2 counts per degree. This value can be changed by increasing the integration (collection) time per reading of the spectrometer. It is clear that there is measurable change in spectrum with change in temperature of the

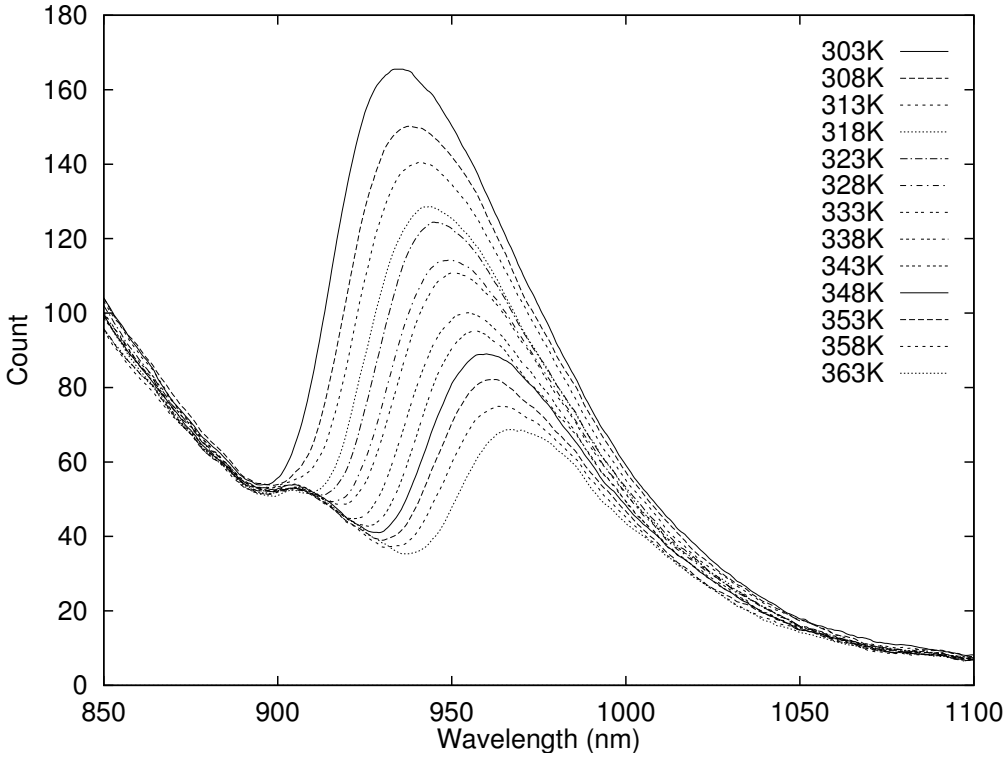


Figure 5.9: Transmission Spectrum of Bulk GaAs

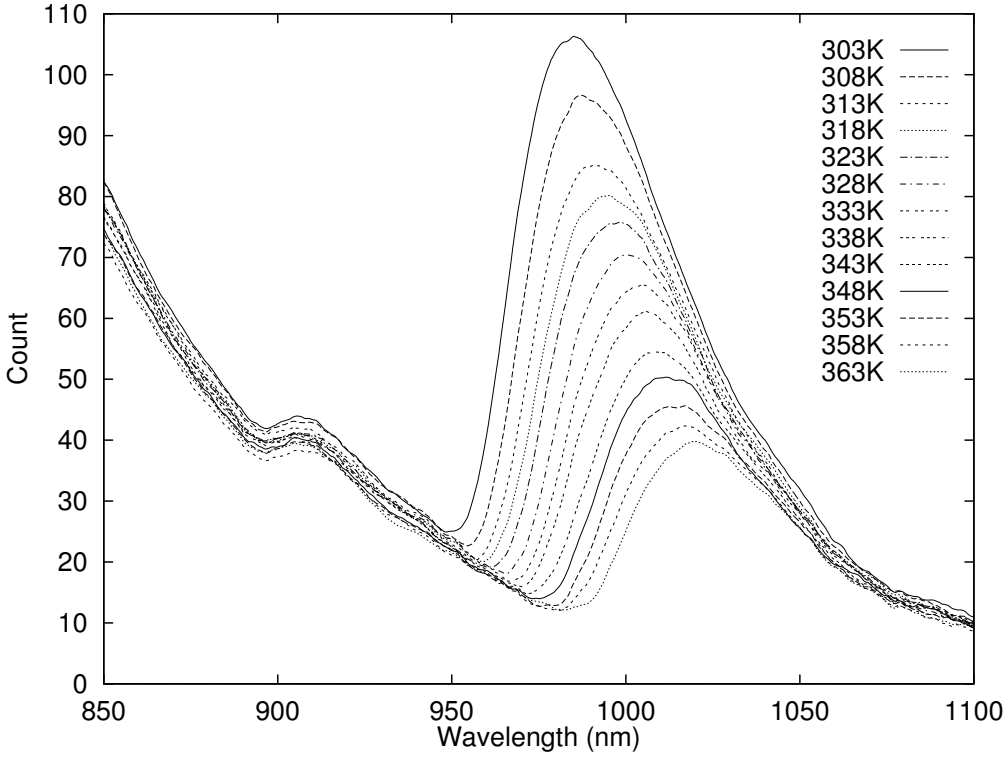


Figure 5.10: Transmission Spectrum of Bulk InP

sample.

### 5.6.5 Conclusion

The electrical experimental results presented represent typical results for the different generations of devices that are repeatable at different times. These results will be used to determine the accuracy of the model developed in preceding chapters and to develop a more robust model.

The preliminary work carried out to determine the feasibility of using optical methods for temperature measurement has shown that the transmission spectra of bulk materials do show changes with change in temperature. As previously mentioned reflectivity, rather than transmission, measurements will be required to determine the temperature of the sample surfaces. This task will be challenging when using thin devices on a thick substrate.

This work was carried out more to determine the capabilities of the digital spectrometer rather than to develop a system for temperature determination. More work is required to develop such a system.

# CHAPTER 6

# MODIFICATION OF DEVICE EQUATIONS

## 6.1 Consideration of Image Force Lowering

Even though no appreciable cooling was observed in the first-generation devices the experimentally obtained $I$-$V$ characteristics are a very useful set of data to compare theoretical models against. Comparison of experimental $I$-$V$ characteristics with those obtained from the original device equations (curves 'A' and 'B' in Fig. 6.2) shows that there is very large disagreement between the two.

One or more assumptions made in section 3.1 are clearly incorrect as seen by the difference in results. One such assumption was that there is no space charge accumulation in the barriers. Because of the small widths of the barriers in these devices the electric field can become very large [53]. This can introduce a space-charge potential that can change the shape of the potential barrier and reduce it's effective height. Scott [54] gives a full treatment of space charge theory and extends the work of Langmuir who summarised and extended previous work dealing with the potential distribution near an electron-emitting plane cathode.

When an electron leaves a cathode surface, it loses momentum due to the potential barrier at the surface and due to the potential field outside the surface. One of the statements made by Scott is that in the region where the image potential is effective (up to about $5\times10^{-8}$ m from the cathode), the space charge potential can be approximated very closely as a linear extrapolation of the initial slope of the space-charge curve. In particular, he notes that image force lowering can approximate space-charge potential very closely if the space-charge potential is linear near the interface between two materials. Image force

lowering refers to when an electron sits outside of a material. It will experience a pseudo force due to a mirror image of the charge with opposite sign appearing to exist inside the material. This pseudo-attractive force will reduce the apparent potential barrier seen by the charge outside the material. As a first approximation, image force lowering is used to approximate space-charge accumulation to determine if space charge effects significantly change the theoretical $I$-$V$ characteristics. Mahan [10] gave a similar treatment to metal-vacuum interfaces to show that the work function can never be reduced below 0.71 eV at room temperature. From a quantum mechanical point of view the solution of effective potential has to be consistent with the electron density distribution. When they are solved in a self-consistent manner, the effect of space-charge is large and gives a minimum work function.

To calculate the image force lowering between two dielectric materials (in this case 2 differing semiconductors), an extension of the derivation for the image force between a dielectric/vacuum interface [55] is made. Referring to Fig. 6.1 the potential of the surface charge is given by

$$\sigma_b = \boldsymbol{P} \cdot \hat{\boldsymbol{n}} = P_z = \epsilon_0 \chi_2 E_z \tag{6.1}$$

$\boldsymbol{P}$ is the polarisation of the dielectric. $E_z$ is the z-component of the total field due to the point charge just inside dielectric 2 at $z = 0$. This field is due in part to $q$ and to the bound charge itself.

$$E_z = -\frac{1}{4\pi\epsilon_1} \cdot \frac{q}{r^2 + d^2} \cdot \cos\theta - \frac{\sigma_b}{2\epsilon_1} = -\frac{1}{4\pi\epsilon_1} \cdot \frac{qd}{\left(r^2 + d^2\right)^{3/2}} - \frac{\sigma_b}{2\epsilon_1} \tag{6.2}$$

The first term is found using Coulomb's Law and the second term is the field due to the bound charge.

Combining Eq. 6.1 and Eq. 6.2 and solving for $\sigma_b$ we get:

$$\sigma_b = -\left[\frac{\chi_2}{2\epsilon_{r1} + \chi_2}\right] \frac{qd}{2\pi \left(r^2 + d^2\right)^{3/2}} \tag{6.3}$$

This expression is exactly the same as the induced charge on an infinite conducting plane under similar circumstances except for the $\chi_2/(2\epsilon_{r1}+\chi_2)$ term. Using this fact and results

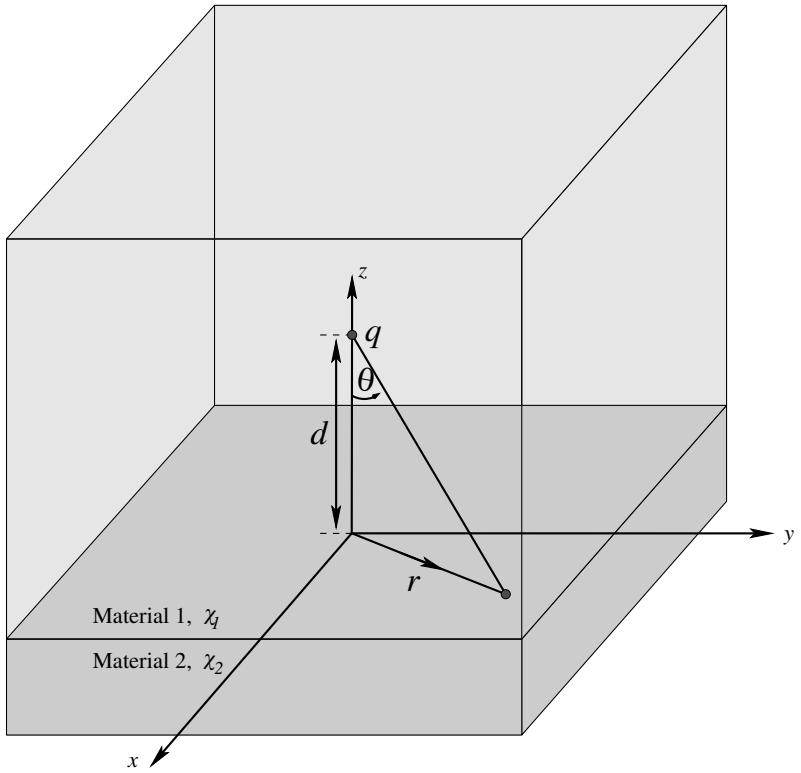


Figure 6.1: Image force of electron between two dielectric materials

given for an infinite conducting plate, the total bound charge is:

$$q_b = -\left[\frac{\chi_2}{2\epsilon_{r1}+\chi_2}\right] q \tag{6.4}$$

and the image force acting on the charge, $q$, is given by:

$$F = -\frac{1}{4\pi\epsilon_0}\left(\frac{\chi_2}{2\epsilon_{r1}+\chi_2}\right)\frac{q^2}{4d^2} = -\frac{1}{4\pi\epsilon_0}\left(\frac{\epsilon_{r2}-1}{2\epsilon_{r1}+\epsilon_{r2}-1}\right)\frac{q^2}{4d^2} \tag{6.5}$$

and the corresponding electric field (using $F = qE$) is:

$$E_z = -\frac{1}{4\pi\epsilon_0}\left(\frac{\epsilon_{r2}-1}{2\epsilon_{r1}+\epsilon_{r2}-1}\right)\frac{q}{4d^2} \tag{6.6}$$

Following the arguments of Scott and using the result above for the image force and electric field it is found that the barrier height of the interface between two dielectrics, $\epsilon_{r1}$ and $\epsilon_{r2}$, is lowered by an amount

$$\Delta\phi = \sqrt{\frac{qE_z}{4\pi\epsilon_0}\cdot\frac{\epsilon_{r2}-1}{2\epsilon_{r2}+\epsilon_{r1}-1}} \tag{6.7}$$

By changing the work function of each barrier in the model developed in chapter 4 by the amount given in Eq. 6.7, image force lowering of the barriers is used to approximate space-charge effects in the model.

### 6.1.1 Analysis of First-Generation Devices with Image Force Lowering of Barrier Heights Considered

The numerical solution given in Chapter 4 was modified so that each barrier height, $\phi_i$, was varied by an amount given in Eq. 6.7. The results are shown in Fig. 6.2.

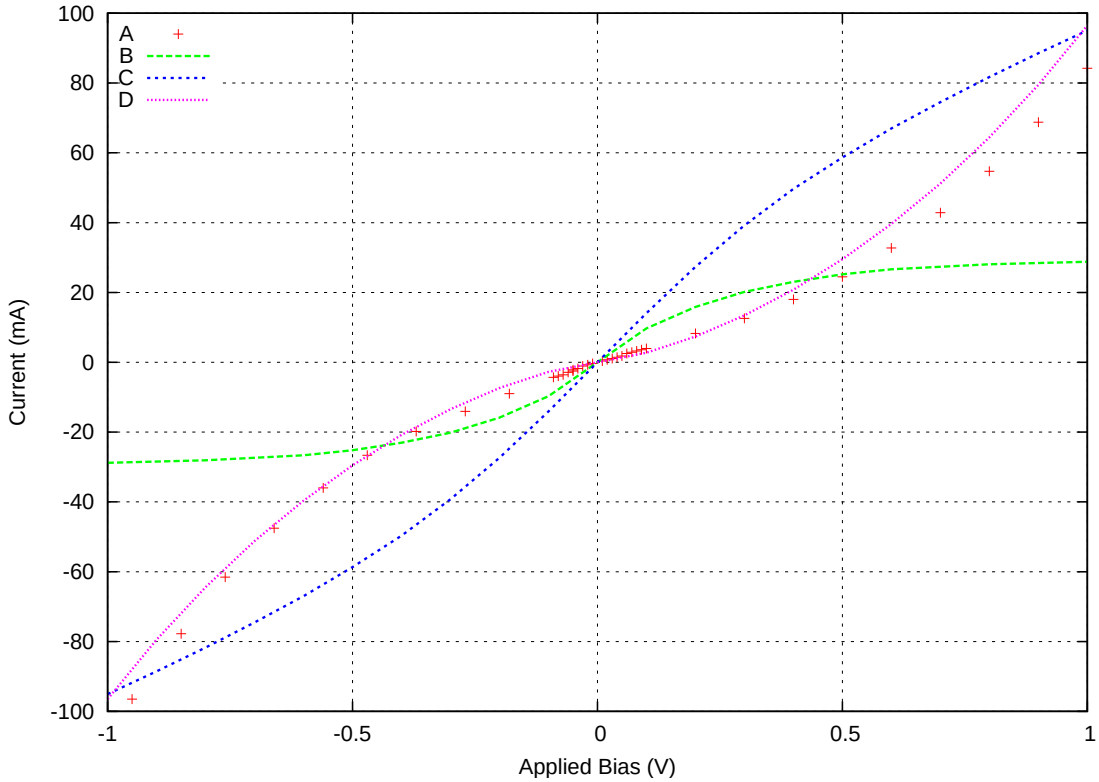


Figure 6.2: $I$-$V$ characteristics of 10-barrier devices: experimental (A) and theoretical without image force lowering (B). Theoretical (C) of same device with image force lowering. Theoretical with exaggerated image force lowering (D) to show that larger barrier-lowering effects bring theory closer to experiment.

Series 'A' is the experimental $I$-$V$ characteristics for a first-generation device with a 6 mm diameter. Series 'B' is the theoretical $I$-$V$ characteristics of the same device generated by the numerical solution presented in Chapter 4. The numerical parameters used are $T_c = T_h = 293$ K, $\phi_i = 0.49$ eV, $m^* = 0.063$, $R_{th} = 9.09 \times 10^{-10}$ Wm$^{-2}$K$^{-1}$, $\epsilon_{r1} = 12.9$ (GaAs) and $\epsilon_{r2} = 12.7012$ ($Al_{0.07}Ga_{0.93}As$). Series 'C' gives the $I$-$V$ characteristics with barrier height lowering due to the image force given by Eq. 6.7. Series 'D' gives

the same characteristics, but with the barrier height lowering increased by a factor of 3. All 3 theoretical curves have been scaled to have the same order of magnitude as the experimental results. The theoretical values are much larger due to the fact that they are for an ‘ideal’ device. The actual magnitude of the results are unimportant at this stage because there is a very large discrepancy between the shape of experimental and theoretical curves. This difference must be addressed first.

Comparing curves ‘A’ and ‘B’ it is immediately obvious that the shape of the two curves is vastly different. Curve ‘B’ saturates at high applied bias. The experimental curve does not saturate.

With the inclusion of barrier lowering due to the image force Curve ‘C’ does show an increase in current at higher applied bias, but still approaches saturation. The barrier lowering due to image force alone does not account for the difference in curves.

To determine whether more substantial barrier lowering can bring the theoretical and experimental curves into closer agreement, curve ‘D’ includes barrier lowering increased by a factor of three. This curve and the experimental curve now are very close in shape, but the amount of barrier lowering cannot be due to image force alone. It is obvious that the space charge potential is no longer adequately approximated by image force lowering and a self-consistent quantum mechanical solution would yield better results.

### 6.1.2 Conclusion

By following the work of Scott [54] inclusion of barrier height lowering due to space charge potential was made by approximating it by image force lowering. By comparing experimental and theoretical curves it was found that this approximation does not adequately describe the real devices.

In order to model the devices more effectively it was decided that a more complete treatment of the devices was necessary. By doing so, inclusion of space charge accumulation

and band bending is possible. The remainder of the chapter will describe a method of using finite differences to create a more complete model of the devices.

## 6.2 Poisson's Equation

The treatment of the thermionic devices in previous chapters has assumed that there is no band bending or space charge accumulation. The inclusion of space-charge into the existing model in section 6.1 showed that the simple treatment given there did not sufficiently explain the experimental $I$-$V$ results obtained.

To further explore these results a first principles approach was taken to analysing the devices. The first step was to obtain a better approximation of the band structure and space-charge through the device. This can be done fairly easily by solving Poisson's equation using a numerical treatment. Although there are presently many computer programs that solve this and other system equations, it was decided to develop a new program that would model the devices in question. This way the exact physics incorporated into the device is known and any modifications or additions can be made in the future. It is expected that the developed program will be extended by future students.

The field of using finite differences to model semiconductor devices is now a very mature field. Many papers have been published on modelling $pn$ junctions, heterostructures and other semiconductor devices [56, 57, 58, 59, 60]. Because of the devices to be considered, a model including degenerate heterostructures is needed. Inclusion of energy transport requires that non-isothermal conditions also be included. Due to time restrictions, energy transport equations are not currently included in the model. In order to allow easy extension to include energy transport in the future, the model developed here does take into account temperature variation through the devices.

As previously mentioned, before electrical transport is modelled in the devices the first step is to generate a better approximation of the energy band diagram of the devices

using Poisson's equation. Poisson's equation is closely related to one of Maxwell's equations [61]:

$$\nabla \cdot \boldsymbol{D} = \rho \tag{6.8}$$

where $\boldsymbol{D} = \epsilon \cdot \boldsymbol{E}$ is the electric flux density and $\boldsymbol{E}$ is the electric field intensity given by:

$$\boldsymbol{E} = -\frac{\partial \psi}{\partial x} \tag{6.9}$$

$\psi$ is the potential at position $x$ in the device.

For a semiconductor the total charge is given by:

$$\rho = q\left(N_D^+ - n + p - N_A^-\right) \tag{6.10}$$

Combining these equations gives us Poisson's equation:

$$\nabla \cdot \boldsymbol{D} = q\left(N_D^+ - n + p - N_A^-\right) \tag{6.11}$$

The terms that appear in Eq. 6.11 are given below.

$N_D^+$ is the number of incompletely ionised donors given by:

$$N_D^+ = \frac{N_D}{1 + g_D \exp\left(\frac{\xi - E_D}{k_B T}\right)} \tag{6.12}$$

where $N_D$ is the donor doping concentration, $g_D$ is the degeneracy factor of the donors and $E_D$ is the energy level of dopants.

$N_A^-$ is the number of incompletely ionised acceptors given by:

$$N_A^- = \frac{N_A}{1 + g_A \exp\left(\frac{E_A - \xi}{k_B T}\right)} \tag{6.13}$$

where $N_A$ is the acceptor doping concentration, $g_A$ is the degeneracy factor of the acceptors and $E_A$ is the energy level of acceptors.

The density of electrons in the material, $n$, is given by:

$$n = N_c \mathcal{F}_{\frac{1}{2}}\left(\frac{\xi_n - E_c}{k_B T}\right) \tag{6.14}$$

where $\xi_n$ is the quasi-Fermi level of electrons. $E_c$ is the conduction band energy measured from the vacuum level and according to Anderson [62] can be found by:

$$E_c = -\psi - \chi \tag{6.15}$$

where $\chi$ is the electron affinity of the material.

The function $\mathcal{F}_{\frac{1}{2}}(x)$ in Eq. 6.14 is a Fermi-Dirac Integral. The family of complete Fermi-Dirac integrals (or simply Fermi integrals) are usually defined as [63]:

$$\mathcal{F}_n(x) = \frac{1}{\Gamma(n+1)} \int_0^\infty \frac{t^n dt}{\exp(t-x)+1}, n > -1 \tag{6.16}$$

where the Gamma function, $\Gamma(x)$, is given by:

$$\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt \tag{6.17}$$

Numerical approximations for Fermi integrals are given by Van Halen and Pulfrey [64] (with a follow-up erratum [65]) and by Aymerich-Humet, Serra-Mestres and Millan [66]

$N_c$ is the bulk electron density of states in the material:

$$N_c = 2\left(\frac{m_n^* k_B T_n}{2\pi\hbar^2}\right)^{\frac{3}{2}} \tag{6.18}$$

The density of holes in the material, $p$, is given by:

$$p = N_v \mathcal{F}_{\frac{1}{2}}\left(\frac{E_v - \xi_p}{k_B T}\right) \tag{6.19}$$

where $\xi_p$ is the quasi-Fermi level of holes and $E_v$ is the valence band energy:

$$E_v = E_c - E_g \tag{6.20}$$

where $E_g$ is the band-gap of the material.

$N_v$ is the bulk hole density of states in the material:

$$N_v = 2\left(\frac{m_p^* k_B T_p}{2\pi\hbar^2}\right)^{\frac{3}{2}} \tag{6.21}$$

For a device under equilibrium conditions with no external applied field the quasi-Fermi level of electrons and holes is constant and equal throughout the whole device:

$$\xi = \xi_n = \xi_p \tag{6.22}$$

So, at any one position in the device there is one equation to solve (Eq. 6.11) and one unknown, $\psi(x)$. Before Poisson's equation is solved throughout the device $\xi$ must be calculated. This can be found simply by solving the charge neutral condition. By first assuming that the potential, $\psi$, is zero throughout the device the chemical potential, $\xi$, can be varied until Eq. 6.10 is zero. By then lining up the chemical potentials throughout the device, the potential at each point can be calculated as a starting point for solving

Poisson's equation.

### 6.2.1 Discretisation of Poisson's Equation

In choosing to model a device by numerical analysis it is necessary to discretise the device due to the fact that finite resources are available. Two common methods are to use finite difference and finite element analysis. The following section will deal with finite differences.

It is common practice to define the system variables of a device at a finite number of points, $N$, as shown in Fig. 6.3 [67]. Auxiliary points, $M$, are defined between these

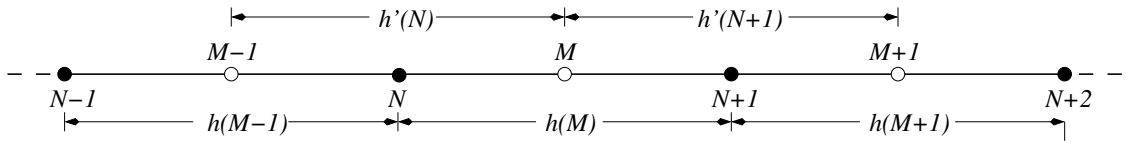


Figure 6.3: Discretisation of Device into Nodes, $N$, and auxiliary points, $M$.

main points. Derivatives of the system variables are calculated at these points.

Eq. 6.11 then is discretised at point $N$ into the following finite difference equation:

$$\frac{D(M) - D(M-1)}{h'(N)} = q\left(N_D^+(N) - n(N) + p(N) - N_A^-(N)\right) \tag{6.23}$$

and the discretised electric flux density at point $M$ is given by:

$$D(M) = -\left(\frac{\epsilon(N+1) + \epsilon(N)}{2}\right) \cdot \left(\frac{\psi(N+1) - \psi(N)}{h(M)}\right) \tag{6.24}$$

The average value of system parameters such as $\epsilon$ is used between main points.

### 6.2.2 Scaling of Variables

In order to increase computational efficiency and decrease redundancy the device equations and variables can be scaled to make them unitless. This also has the advantage of reducing the range that terms can take. They are scaled by using the scheme proposed by De Mari [68] and are shown in Table 6.1.

Table 6.1: Scaling Variables

| **Quantity / Parameter** | **Description** | **Normalisation Factor** |
|---|---|---|
| $\psi$ | electrostatic potential | $V_t = k_B T/q$ |
| $N_c, N_v$ | electron / hole density of states | $n_{i0}$ |
| $N_A, N_D$ | acceptor / donor concentrations | $n_{i0}$ |
| $x$ | position coordinate | $x_0 = \sqrt{\epsilon V_t / q n_{i0}}$ |
| $E_g$ | band-gap | $qV_t$ |
| $\chi$ | electron affinity | $qV_t$ |
| $E_A$, $E_D$ | acceptor / donor energy levels | $qV_t$ |
| $D_n, D_p$ | electron / hole diffusion constants | $D_0 = 10^{-4}\text{m}^2\text{s}^{-1}$ |
| $\mu_n, \mu_p$ | electron / hole mobilities | $D_0/V_t$ |
| $\epsilon$ | permittivity | $\epsilon_0$ |
| $t_n$, $t_p$ | Shockley-Read-Hall lifetimes | $x_0^2/D_0$ |
| $C_n$, $C_p$ | Auger coefficients | $D_0/(x_0^2 n_{i0}^2)$ |
| $C_{opt}$ | Optical Capture-Emission Rate | $D_0/(x_0^2 n_{i0})$ |
| $\alpha_n^{max}$, $\alpha_p^{max}$ | Impact Ionisation coefficients | $1/x_0$ |
| $E_n^{crit}$, $E_p^{crit}$ | Impact Ionisation critical fields | $x_0/V_t$ |
| $J_0$ | Carrier Current Density | $J_0 = qD_0 n_{i0}/x_0$ |
| $A^*$ | Richardson's Constant | $J_0$ |

### 6.2.3 Band Discontinuities

Conduction- and valence- band discontinuities can generally be calculated using Anderson's electron-affinity rule [69]. This rule states that at a heterojunction interface the conduction band offset is given by:

$$\Delta E_c = \chi_2 - \chi_1 \tag{6.25}$$

where $\chi_1$ and $\chi_2$ are the electron affinities of the two materials. The valence-band offset is given by:

$$\Delta E_v = \Delta E_g - \Delta E_c \tag{6.26}$$

The widely accepted band-offset ratio for GaAs/AlGaAs, $\Delta E_c : \Delta E_v = 85 : 15$, deduced by Dingle et al. [70] has been replaced, after many sophisticated measurements over recent years, by a value closer to $\Delta E_c : \Delta E_v = 63 : 35$.

Table 6.2: $Al_xGa_{1-x}As$ Material Parameters

| **Name** | **Value** | | | **Unit** |
|---|---|---|---|---|
| $m_n$ | $0.063 + 0.083x$<br>$0.85 - 0.14x$ | $x < 0.45$<br>$0.45 < x < 1$ | [42] | $m_0$ |
| $m_p$ | $0.51 + 0.25x$ | | [42] | $m_0$ |
| $\mu_n$ | $8000 - 22000x + 10000x^2$<br>$-255 + 1160x - 720x^2$ | $x < 0.45$<br>$0.45 < x < 1$ | [42] | $cm^2/V{\cdot}s$ |
| $\mu_p$ | $370 - 970x + 740x^2$ | | [42] | $cm^2/V{\cdot}s$ |
| $E_g$ | $1.519 + 1.155x + 0.37x^2 - 5.41 \cdot 10^{-4} \frac{T^2}{T+204}$<br>$1.981 + 0.124x + 0.144x^2 - 4.6 \cdot 10^{-4} \frac{T^2}{T+204}$ | $x < 0.45$<br>$0.45 < x < 1$ | [42] | eV |
| $\chi$ | $4.07 - 0.79x$<br>$3.595 + 0.335x - 0.143x^2$ | $x < 0.45$<br>$0.45 < x < 1$ | Sec. 6.2.3 | eV |
| $\epsilon_r$ | $12.90 - 2.84x$ | | [42] | $\epsilon_0$ |
| $g_D$ | 2 | | | – |
| $g_A$ | 4 | | | – |
| $E_D$ | $5.2 + 7.9x + 7.1x^2$ | | [42] | meV |
| $E_A$ | $30.7 + 495x^{2.78}$ | | [42] | meV |
| $t_n$ | $\sim 1 \times 10^{-8}$ | | [42] | s |
| $t_p$ | $\sim 1 \times 10^{-8}$ | | [42] | s |
| $C_n$ | $\sim 1.2 \times 10^{-43}$ | | [42] | $m^6/s$ |
| $C_p$ | $\sim 8.5 \times 10^{-43}$ | | [42] | $m^6/s$ |
| $C_{opt}$ | $\sim 1 \times 10^{-16}$ | | [42] | $m^6/s$ |
| $E_n^{crit}$ | $\sim 6.31 \times 10^7$ | | [42] | V/m |
| $E_p^{crit}$ | $\sim 7.22 \times 10^7$ | | [42] | V/m |
| $\alpha_n^{max}$ | $\sim 1.81 \times 10^7$ | | [42] | $m^{-1}$ |
| $\alpha_p^{max}$ | $\sim 3.05 \times 10^7$ | | [42] | $m^{-1}$ |
| $\beta_{n,p}$ | 2 | | [42] | – |

The commonly published value for the electron affinity of AlGaAs:

$$\chi = 4.07 - 1.1x, \qquad 0 < x \leq 0.45$$

$$\chi = 3.64 - 0.14x, \quad 0.45 < x < 1$$

does not result in the correct band-offset ratios. Because only differences in electron

affinity appear in the device equations we can effectively replace the published value with one that does give the correct ratio. The reason that this is done is because Anderson's rule (using only electron affinities) does not correctly predict the band-offset for some materials, such as GaAs/AlGaAs. By modifying the expression for electron affinity we can then use Anderson's rule to calculate band structure.

Currently accepted values for the conduction band offset of GaAs/AlGaAs are [42]:

$$\begin{aligned} \Delta E_c &= 0.79x, & 0 < x \leq 0.45 \\ \Delta E_c &= 0.475 - 0.335x + 0.143x^2, & 0.45 < x < 1 \end{aligned}$$

Using this we can deduce the electron affinity that, using Anderson's rule, will give the correct band-offsets and band-offset ratio between GaAs and $Al_xGa_{1-x}As$ at 300 K:

$$\chi = 4.07 - 0.79x, \qquad 0 < x \leq 0.45 \tag{6.27}$$

$$\chi = 3.595 + 0.335x - 0.143x^2, \qquad 0.45 < x < 1 \tag{6.28}$$

Due to the uncertainty in these values, and the little effect that temperature has on the change in band-gap, no temperature dependence is needed and these expressions are assumed to hold for all temperatures of interest (near 300 K).

### 6.2.4 Boundary Conditions

The potential of the left device contact is always assumed to be zero. The right contact is assumed to be at a potential equal to the built-in potential of the device $V_{bi}$. This value is found by solving the charge-neutral condition described earlier. The chemical potential throughout the device is also found during this step and is held constant.

### 6.2.5 Newton-Raphson Method of Solving Poisson's Equation

The now familiar technique of using the Newton-Raphson method is once again employed to solve Poisson's equation. For an $N$ node system the Jacobian Matrix is calculated numerically by taking the partial derivative of Poisson's equation at each node ($i = 2, 3, \ldots N-1$) with respect to the potential at each node. The Jacobian is used to

update the potential of each node at each iteration until the relative error is less than a preset tolerance.

#### 6.2.5.1 Optimisation of Newton's Method

By inspecting Eq. 6.23 it can be seen that Poisson's equation at each node is only dependent on the potential of three nodes in the device according to:

$$\begin{aligned} \frac{\partial D_i}{\partial \psi_j} &\neq 0, \quad j = i-1, i, i+1; \\ &= 0, \quad \text{otherwise.} \end{aligned} \tag{6.29}$$

For a 200-node device the Jacobian would be a sparse $198 \times 198$ matrix with all entries zero except for the diagonal and sub- and super- diagonals. This means that 39204 separate calculations are needed to populate the Jacobian. The inversion of such a matrix is computationally expensive and is a waste of resources due to the many zero entries.

By using the information provided by Eq. 6.29 only $3 \times 198 = 594$ calculations are needed. Such a *tridiagonal* matrix can be inverted very efficiently by using special algorithms for such a purpose [46]. Indeed, experimentation showed that calculation and inversion of the $198 \times 198$ matrix could take up to 5 minutes per iteration. By employing tridiagonal techniques, the same iteration can be achieved in a matter of milliseconds. Also the accuracy and stability of such a technique is improved because of the absence of many more redundant calculations that can introduce a whole gamut of numerical problems such as round-off error.

## 6.3 Generating the Energy Band Structure of a First-Generation device

Simple text files with a set format are used to define the structure of devices to be modelled by the program, called *TISim*. An example of a device is given in Fig. 6.4. This input file defines the structure of a first-generation device as shown in Fig. 5.1. The substrate and top buffer layer have been excluded because they are not part of the active region of the device. 100 nm $n^+$-GaAs regions have been included on the top and bottom of the device instead as these regions set the chemical potential throughout the device.

```
Mat=AlGaAs x=0.0  Nd=2e24 Width=100 Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50  Step=5
Mat=AlGaAs x=0.0  Nd=2e24 Width=100 Step=5
```

Figure 6.4: Example Listing of a first-generation thermionic device.

#### 6.3.0.2 Input File Format

The example input file instructs *TISim* that the device is made of $\mathrm{Al}_x\mathrm{Ga}_{1-x}\mathrm{As}$ alloy with the aluminium concentration alternating between 0 and 7%. The C++ source file 'Material.cpp' can be modified to allow any material to be used. The next column specifies the donor and/or acceptor concentrations in units of $\mathrm{m}^{-3}$. The width is then specified in nm with the distance step between nodes also specified in nm. Columns can appear in any order, no spaces are permitted between the parameter name, the '=' sign and the parameter value and lines beginning with '#' are ignored as comments. Columns must be separated by white space (either one or more spaces or tabs).

### 6.3.1 Energy Band Diagram of a First-Generation Device

When solved for this device, Poisson's equation gives the energy band diagram shown in Fig. 6.5. Both Fermi-Dirac and Maxwell-Boltzmann statistics were used and are included

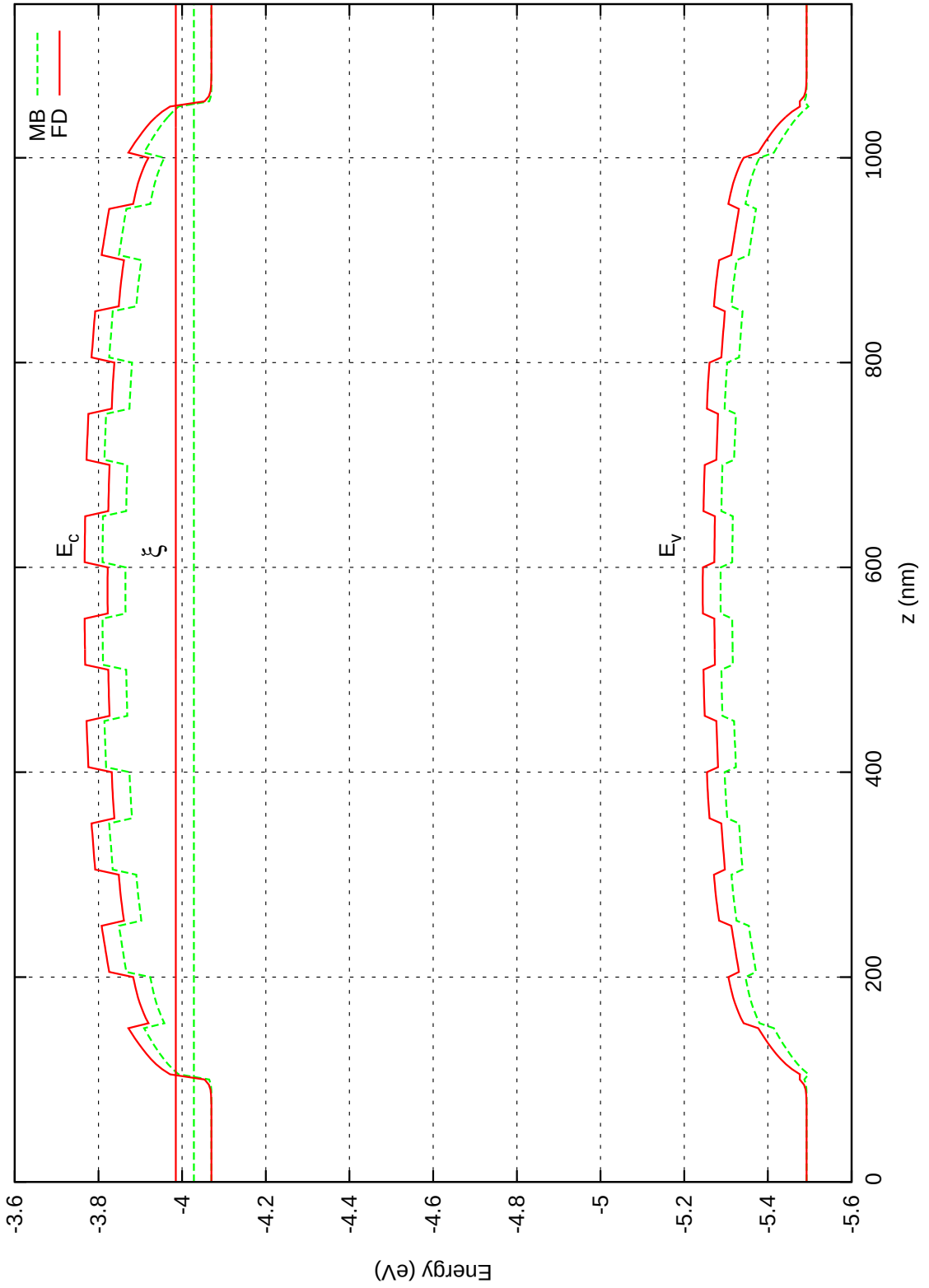


Figure 6.5: Band Diagram of 10-barrier heterostructure device using Maxwell-Boltzmann (MB) and Fermi-Dirac (FD) statistics

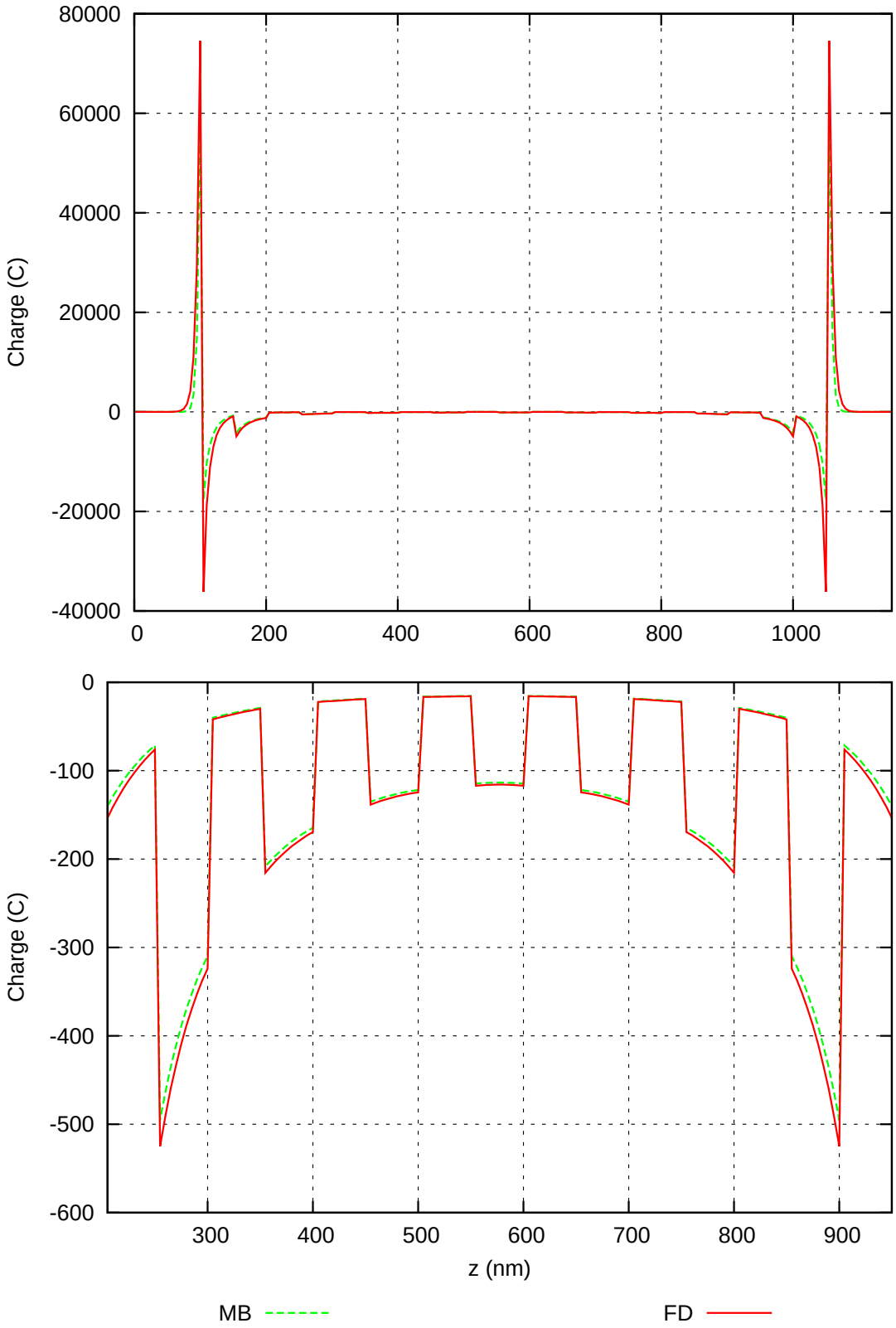


Figure 6.6: Space charge of 10-barrier heterostructure device using Maxwell-Boltzmann (MB) and Fermi-Dirac (FD) statistics

in the figure for comparison. Because of the high doping concentration there is quite a large discrepancy between the two. It is immediately obvious that there is considerable bending of the conduction and valence bands thus bringing the accuracy of the simple model formed in chapter Richardson's equation into question.

Fig. 6.6 shows the space-charge accumulation in the device. The bottom plot shows an expanded view of the device between 205 nm and 950 nm. In this region there is a quasi-linear variation of space charge accumulation (and corresponding potential) near the heterostructure interfaces. As discussed in Section 6.1 image force lowering should be a valid approximation for space-charge accumulation. But, referring to the top plot, there are regions at both ends of the device in which there is clearly a non-linear variation in space-charge. In these regions the image-force approximation will be incorrect. This most likely accounts for the inability of the simple model including image-force lowering to model the $I$-$V$ characteristics of the devices properly.

By using a more detailed model such as this it is hoped that more accurate $I$-$V$ characteristics will be available. In order for carrier transport to be included in the model the next section discusses incorporation of Boltzmann's Transport Equation.

## 6.4 Boltzmann Transport Equation

Poisson's equation calculates the energy diagram and charge concentration of a device under equilibrium conditions. When an external bias is applied to the device the quasi-Fermi levels are no longer constant and, in general, no longer equal. Two new equations are required in order to calculate how carriers move under an external field and thus how the probability of average energy (quasi-Fermi levels) change.

These two equations can both be derived from the *Boltzmann Transport Equation*. This equation calculates how the two carrier concentrations (electrons and holes) vary from the equilibrium distribution function with time when under the influence of external forces.

The validity of the transport equation is dependent on the sample size, $L$, $L \gg \lambda \gg a$.

When $L \sim a$ we need to include quantum corrections. When $L \sim \lambda$, randomising scattering processes are absent and the carriers are ballistic [71].

Starting with the first order approximation of the Boltzmann Transport Equation [37]:

$$\frac{\partial f}{\partial \boldsymbol{r}}\frac{\partial \boldsymbol{r}}{\partial t} + \frac{\partial f}{\partial \boldsymbol{k}}\frac{\partial \boldsymbol{k}}{\partial t} + \frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{coll} \tag{6.30}$$

and using

$$\frac{\hbar d\boldsymbol{k}}{dt} = -q\left(\boldsymbol{\mathcal{E}} + \boldsymbol{v} \times \boldsymbol{\mathcal{B}}\right) \tag{6.31}$$

we get the equation of motion:

$$\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f + \frac{q}{\hbar}\left(\boldsymbol{\mathcal{E}} + \boldsymbol{v} \times \boldsymbol{\mathcal{B}}\right) \cdot \nabla_{\boldsymbol{k}} f + \frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{coll} \tag{6.32}$$

The collision term on the right is difficult to calculate and the *relaxation time approximation* is often used:

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\frac{f - f_0}{\tau(\boldsymbol{k})} \tag{6.33}$$

This approximation assumes that if the carrier distribution function, $f$, differs from the equilibrium distribution it will return to equilibrium in characteristic time $\tau(\boldsymbol{k})$.

For steady-state conditions in which the time dependence of $f$ on $t$ is ignored, we arrive at:

$$\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f + \frac{q}{\hbar}\left(\boldsymbol{\mathcal{E}} + \boldsymbol{v} \times \boldsymbol{\mathcal{B}}\right) \cdot \nabla_{\boldsymbol{k}} f = -\frac{f - f_0}{\tau(\boldsymbol{k})} \tag{6.34}$$

In the devices studied the distribution function varies with distance so $\nabla_{\boldsymbol{r}} f$ is non-zero. We want to look at the devices under an external electric field, $\boldsymbol{\mathcal{E}}$, with no external magnetic field, $\boldsymbol{\mathcal{B}}$. This then gives us the general equation of motion of carriers that can be used to develop a more useful model:

$$\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f + \frac{q}{\hbar}\boldsymbol{\mathcal{E}} \cdot \nabla_{\boldsymbol{k}} f = -\frac{f - f_0}{\tau(\boldsymbol{k})} \tag{6.35}$$

If we further assume that the deviation of $f$ from the equilibrium distribution $f_0$ is small, we can replace $f$ with $f_0$ on the left hand side of Eq. 6.35. This gives the approximate form of the distribution function to be used:

$$\boxed{f = f_0 - \tau(\boldsymbol{k})\left(\boldsymbol{v} \cdot \nabla_{\boldsymbol{r}} f_0 + \frac{q}{\hbar}\boldsymbol{\mathcal{E}} \cdot \nabla_{\boldsymbol{k}} f_0\right)} \tag{6.36}$$

## 6.5 Derivation of Carrier Current Densities due to drift-diffusion

The devices will be modelled by a one-dimensional treatment. That is, the devices are anisotropic in the $x$-direction and isotropic in both the $y$- and $z$- directions. Eq. 6.36 can be used to calculate the current density in the $x$-direction, $J_x$. The equilibrium distribution function is given by the Fermi-Dirac occupation function:

$$f_{n0}(E) = \frac{1}{1+\exp\left(\frac{E_n-\xi_n}{k_BT_n}\right)} \quad \text{for electrons,} \tag{6.37}$$

$$f_{p0}(E) = \frac{1}{1+\exp\left(\frac{\xi_p-E_p}{k_BT_p}\right)} \quad \text{for holes.} \tag{6.38}$$

Approximating these by their corresponding Maxwell-Boltzmann functions simplifies calculations but because the devices in question have highly doped regions this approximation may diverge from the correct solution.

If we assume that the kinetic energy, $E_k$, of $n$- and $p$- type carriers is related to the conduction, $E_c$, and valence, $E_v$, bands respectively, by a simple spherical approximation [38] we have the total energy of the carriers:

$$E_n(\boldsymbol{k}) = E_c + E_k = E_c + \frac{1}{2}m_n\boldsymbol{v}^2 = E_c + \frac{\hbar^2\boldsymbol{k}^2}{2m_n} \tag{6.39}$$

$$E_p(\boldsymbol{k}) = E_v - E_k = E_v - \frac{1}{2}m_p\boldsymbol{v}^2 = E_v - \frac{\hbar^2\boldsymbol{k}^2}{2m_p} \tag{6.40}$$

we have the relationship:

$$\nabla_{\boldsymbol{k}}E_{n,p} = \frac{\hbar^2\boldsymbol{k}}{m_{n,p}} = \hbar\boldsymbol{v}_{n,p} \tag{6.41}$$

### 6.5.1 Derivation of Drift-Diffusion Current for a Degenerate Non-Isothermal Heterostructure

The following derivation calculates the current density for electrons. For the rest of the derivation treat $f_0 = f_{n0}$.

We begin with the general equation for carrier current density and substitute the approximate form of the distribution function given be Eq. 6.36:

$$\begin{aligned} J_n &= -\frac{q}{4\pi^3}\int \boldsymbol{v} f d^3\boldsymbol{k} \\ &= -\frac{q}{4\pi^3}\int \boldsymbol{v} f_0 d^3\boldsymbol{k} + \frac{q}{4\pi^3}\int \tau(\boldsymbol{k})\boldsymbol{v}\left(\frac{q}{\hbar}\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{k}}f_0 + \boldsymbol{v}\cdot\nabla_r f_0\right)d^3\boldsymbol{k} \end{aligned}$$

The first term vanishes as it involves a symmetrical integral of an odd function, leaving:

$$J_n = \underbrace{\frac{q^2}{4\pi^3\hbar}\int \tau(\boldsymbol{k})\boldsymbol{v}\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{k}} f_0 d^3\boldsymbol{k}}_{\text{Drift Current}} + \underbrace{\frac{q}{4\pi^3}\int \tau(\boldsymbol{k})\boldsymbol{v}\boldsymbol{v}\cdot\nabla_r f_0 d^3\boldsymbol{k}}_{\text{Diffusion Current}} \quad (6.42)$$

#### 6.5.1.1 Calculation of Drift Current

$$\boldsymbol{\mathcal{E}} = -\frac{\nabla_{\boldsymbol{r}} E_c}{q} = -\frac{\nabla_x E_c \hat{\boldsymbol{x}}}{q} = \mathcal{E}_x\hat{\boldsymbol{x}}$$

because $(\frac{\partial E_c}{\partial y}, \frac{\partial E_c}{\partial z}) = 0$. Therefore only the $x$-component of $\nabla_{\boldsymbol{k}} f_0$ is needed. For the 1-dimensional case:

$$\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{k}} f_0 = -\frac{\nabla_x E_c}{q}\frac{\partial f_0}{\partial k_x}$$

Then the drift current becomes:

$$J_x^{drift} = \frac{q^2\mathcal{E}_x}{4\pi^3 m_n}\int \tau(\boldsymbol{k})k_x\frac{\partial f_0}{\partial k_x}d^3\boldsymbol{k}$$

By defining the mobility as:

$$\begin{aligned}\mu &= \frac{q}{4\pi^3 n m_n}\int \tau(\boldsymbol{k})k_x\frac{\partial f_0}{\partial k_x}d^3\boldsymbol{k} \\ &= \frac{q}{m_n}\frac{\int \tau(\boldsymbol{k})k_x\frac{\partial f_0}{\partial k_x}d^3\boldsymbol{k}}{4\pi^3 n} \\ &= \frac{q}{m_n}\frac{\int \tau(\boldsymbol{k})k_x\frac{\partial f_0}{\partial k_x}d^3\boldsymbol{k}}{\int f_0 d^3\boldsymbol{k}}\end{aligned} \quad (6.43)$$

we arrive at the Drude form of the drift current [37]:

$$\boxed{J_x^{drift} = q\mu n\mathcal{E}_x = -\mu n\nabla E_c} \quad (6.44)$$

#### 6.5.1.2 Calculation of Diffusion Current

$$\boldsymbol{v}\cdot\nabla_r f_0 = \frac{\hbar k_x}{m_n}\frac{\partial f_0}{\partial x}$$

because $(\frac{\partial f_0}{\partial y}, \frac{\partial f_0}{\partial z}) = 0$.

$$J_x^{diffusion} = \frac{q\hbar^2}{4\pi^3 m_n^2}\int \tau(\boldsymbol{k})k_x^2\frac{\partial f_0}{\partial x}d^3\boldsymbol{k}$$

Using the following relation:

$$\frac{\partial f_0}{\partial x} = \frac{\partial f_0}{\partial E_c}\nabla E_c + \frac{\partial f_0}{\partial \xi_n}\nabla\xi_n + \frac{\partial f_0}{\partial E_k}\nabla E_k + \frac{\partial f_0}{\partial T}\nabla T$$

and following the treatment given by Stratton [72], we have:

$$\frac{\partial f_0}{\partial x} = -\frac{\partial f_0}{\partial E_k}\left[-\nabla E_c + \nabla \xi_n + (E_c - \xi_n + E_k)\frac{\nabla T}{T}\right] \tag{6.45}$$

In order to eliminate $\nabla \xi_n$ we use the expression for $n$ (Eq. 6.14) and the following property for Fermi integrals:

$$\frac{d\mathcal{F}_n(x)}{dx} = \mathcal{F}_{n-1}(x) \tag{6.46}$$

to obtain:

$$\frac{1}{n}\frac{\partial n}{\partial x} = \frac{3}{2}\frac{\nabla m_n}{m_n} + \frac{3}{2}\frac{\nabla T}{T} + \frac{\mathcal{F}_{-\frac{1}{2}}}{\mathcal{F}_{\frac{1}{2}}}\frac{1}{k_B T}\left(\nabla \xi_n - \nabla E_c + (E_c - \xi_n)\frac{\nabla T}{T}\right)$$

Substitution into Eq. 6.45 gives the result:

$$\frac{\partial f_0}{\partial x} = -\frac{\partial f_0}{\partial E_k}\left[\Lambda_c k_B T \mathcal{Z} + E_k \frac{\nabla T}{T}\right]$$

where

$$\Lambda_c = \frac{\mathcal{F}_{\frac{1}{2}}(\eta_c)}{\mathcal{F}_{-\frac{1}{2}}(\eta_c)}$$

$$\mathcal{Z} = \frac{\nabla n}{n} - \frac{3}{2}\frac{\nabla m_n}{m_n} - \frac{3}{2}\frac{\nabla T}{T}$$

We now have an intermediate form for electron current density due to diffusion:

$$J_x^{diffusion} = -\underbrace{\frac{q}{4\pi^3 m_n}\Lambda_c k_B T \mathcal{Z}\int \tau(\boldsymbol{k}) k_x \frac{\partial f_0}{\partial k_x} d^3\boldsymbol{k}}_{A} - \underbrace{\frac{q}{4\pi^3 m_n}\frac{\nabla T}{T}\int \tau(\boldsymbol{k}) k_x E_k \frac{\partial f_0}{\partial k_x} d^3\boldsymbol{k}}_{B}$$

To solve A, compare to expression for mobility (Eq. 6.43) to give:

$$A = n\mu_n \Lambda_c k_B T \mathcal{Z}$$

Solving B, once again compare to mobility to get:

$$B = n\mu_n \frac{\nabla T}{T}\frac{\int \tau(\boldsymbol{k}) k_x E_k \frac{\partial f_0}{\partial k_x} d^3\boldsymbol{k}}{\int \tau(\boldsymbol{k}) k_x \frac{\partial f_0}{\partial k_x} d^3\boldsymbol{k}}$$

Up to this point, the exact form for the momentum relaxation time, $\tau(\boldsymbol{k})$, has been inconsequential because of the definition of mobility. Depending on the type of scattering mechanisms it is usual to have $\tau$ be dependent on the kinetic energy of the carrier to some

power [38]:

$$\tau(\boldsymbol{k}) = \tau(E_k) = \tau_0 E_k^{-S} \tag{6.47}$$

Then, applying the same techniques as used by Seeger [36] in calculating the momentum relaxation time we arrive at a solution for B, also given by Apanovich et al. [73]:

$$B = n\mu_n k_B \nabla T \left(\frac{5}{2} + S\right) \Omega_c \tag{6.48}$$

where

$$\Omega_c = \frac{\mathcal{F}_{S+\frac{3}{2}}(\eta_c)}{\mathcal{F}_{S+\frac{1}{2}}(\eta_c)} \tag{6.49}$$

Combining A and B gives an expression for the current density due to diffusion:

$$\boxed{J_x^{diffusion} = -n\mu_n \Lambda_c k_B T \mathcal{Z} - n\mu_n k_B \nabla T \left(\frac{5}{2} + S\right) \Omega_c} \tag{6.50}$$

Combining the expressions for drift and diffusion current, we arrive at an expression for current due to drift-diffusion that is valid for a non-isothermal, degenerate heterostructure:

$$\begin{aligned} J_n = & - \mu_n \Lambda_c k_B T \nabla n - \mu n \nabla E_c \\ & + \frac{3}{2}\mu_n n \Lambda_c k_B T \frac{\nabla m_n}{m_n} + \frac{3}{2}\mu_n n \Lambda_c k_B \nabla T \\ & - \left(\frac{5}{2} + S\right) \mu_n n \Omega_c k_B \nabla T \end{aligned} \tag{6.51}$$

This form of the equation is necessary so as to be able to compare to the form presented by Smith and Rohatgi [74]. They present a nearly identical expression except that they assume that $\tau(E) = \tau_0$ and is independent of carrier energy. Eq. 6.51 presents a more general case in which different scattering mechanisms can be included by modifying $\tau_0$ and $S$. The reason that this comparison is necessary is because they give discrete forms of the equation to be used in numerical analysis. This topic will be covered further in section 6.8.

### 6.6 Current Due to Thermionic Emission

When there is an abrupt change in band structure due to a heterojunction interface, thermionic emission rather than drift-diffusion dominates current density. Referring to

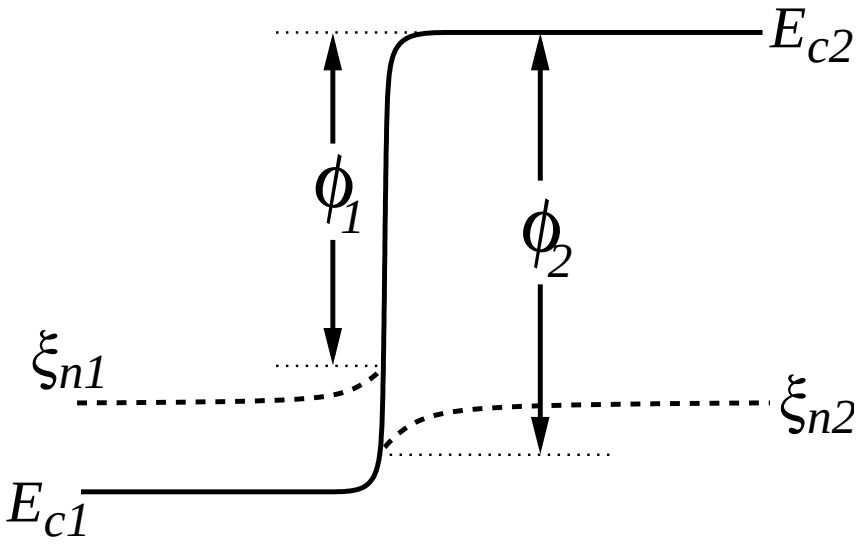


Figure 6.7: Band Discontinuity at abrupt heterojunction

Fig. 6.7 and assuming Maxwell-Boltzmann statistics (see section 2.3 for valid conditions), the current due to thermionic emission is only slightly different to the forms previously derived. The net current due to electrons moving to the right at the interface is given by:

$$J_n^{therm} = -A_n^* T_n^2 \exp\left(-\frac{\phi_1}{k_B T_n}\right) + A_n^* T_n^2 \exp\left(-\frac{\phi_2}{k_B T_n}\right) \tag{6.52}$$

Because of the choice a variables ($\psi$, $n$ and $p$) a more useful form of net thermionic current is given by [75, 76]:

$$J_n^{therm} = -A_n^* T_n^2 \frac{n}{N_c} \exp\left(-\frac{\Delta E_c}{k_B T_n}\right) + A_n^* T_n^2 \frac{n}{N_c} \tag{6.53}$$

This form increases execution time because $n$ and $p$ are usually already calculated at each iteration.

Similarly the thermionic current due to holes is given by:

$$J_p^{therm} = A_p^* T_p^2 \frac{p}{N_v} \exp\left(-\frac{\Delta E_v}{k_B T_n}\right) - A_p^* T_p^2 \frac{p}{N_v} \tag{6.54}$$

These equation are valid for when $E_g^+ > E_g^-$. They must be modified appropriately for when $E_g^+ < E_g^-$.

## 6.7 Steady-State Conditions

### 6.7.1 Continuity Equations

The three main equations that must be solved at each node in the device are continuity equations that relate the electric field to the carrier concentrations and relate the rate of change of carrier concentrations to the net recombination-generation and current density:

$$-\nabla \cdot (\epsilon \nabla \cdot \psi) = \rho \tag{6.55}$$

$$\frac{\partial n}{\partial t} = -(R - G) - \nabla \cdot J_n \tag{6.56}$$

$$\frac{\partial p}{\partial t} = -(R - G) + \nabla \cdot J_p \tag{6.57}$$

The first equation is Poisson's equation derived earlier and the second two are the current continuity equations. By considering only steady-state conditions with a DC applied bias, the two time-dependent variables, $\frac{\partial n,p}{\partial t}$, become zero. The two current continuity equations then state that the current throughout the device must be constant except for net recombination-generation of carriers.

### 6.7.2 Carrier Recombination and Generation Rates

The terms $R$ and $G$ in Eq. 6.56 and Eq. 6.57 refer to recombination and generation of carriers, respectively. The relative significance of generation-recombination effects depends on the particular device in question [77]. For thoroughness and consistency a number of recombination-generation mechanisms are included but it will be shown *a posteriori* that they have little or no effect on the current equations in the devices to be considered.

The net generation-recombination in the model can be considered as the result of several distinct processes: thermal recombination due to phonon transitions can be modelled using the Shockley-Read-Hall model; three-particle interactions are modelled using the Auger model; and photon (optical) transitions. The net recombination rate can then be given as:

$$R_{net} = R_{SRH} + R_{Aug} + R_{opt} \tag{6.58}$$

The net generation of carriers is modelled by considering only one mechanism, namely

impact ionisation:

$$G_{net} = G_{impact} \tag{6.59}$$

#### 6.7.2.1 Shockley-Read-Hall Recombination

The Shockley-Read-Hall model considers recombination due to phonon transitions occurring as the result of traps. This mechanism involves a trap center in the energy bandgap with a two-step capture and emission process. It is common practice to consider only a single energy level trap in the middle of the forbidden band and can then be given by:

$$R_{SRH} = \frac{pn - n_i^2}{t_n\left(p + n_i\right) + t_p\left(n + n_i\right)} \tag{6.60}$$

where $n_i$, the intrinsic carrier concentration, is found from the fact that for a device under equilibrium conditions:

$$n_i^2 = np = N_c N_v \mathcal{F}_{\frac{1}{2}}\left(\eta_n\right) \mathcal{F}_{\frac{1}{2}}\left(\eta_p\right) \tag{6.61}$$

For a nondegenerate semiconductor this reduces to the usual expression:

$$n_i^2 = N_c N_v \exp\left(-\frac{E_g}{k_B T}\right) \tag{6.62}$$

#### 6.7.2.2 Auger Recombination

Auger recombination is characterised by an electron-hole pair recombining without the aid of a trapping center, where energy is given to another electron or hole:

$$R_{Aug} = (pn - n_i^2)(nC_n + pC_p) \tag{6.63}$$

$C_n$ and $C_p$ are Auger coefficients that are specified for a given material. This type of recombination becomes important in regions where there are both high electron and hole concentrations. In the devices considered here, only one type of doping species is considered so it is expected that this type of recombination will be of little importance.

#### 6.7.2.3 Optical Recombination

This process involves the direct transition of carriers between the valence and conduction bands by gaining energy from or emitting energy as photons. Thus it is only significant

in direct band-gap materials:

$$R_{opt} = C_{opt}(pn - n_i^2) \tag{6.64}$$

$C_{opt}$ is the optical capture-emission rate given for a particular material.

#### 6.7.2.4 Impact Ionisation

Impact ionisation can be considered as the reverse of the Auger recombination. Carrier pairs are generated by carriers moving directly across the band-gap:

$$G_{impact} = \frac{\alpha_n |J_n| + \alpha_p |J_p|}{q} \tag{6.65}$$

where $\alpha_n$ and $\alpha_p$ are the ionisation coefficients. These coefficients depend on the electric field parallel to the direction of current flow:

$$\begin{aligned} \alpha_n &= \alpha_n^{max} \exp\left[-\left(\frac{E_n^{crit}}{E}\right)^{\beta_n}\right] \\ \alpha_p &= \alpha_p^{max} \exp\left[-\left(\frac{E_p^{crit}}{E}\right)^{\beta_p}\right] \end{aligned} \tag{6.66}$$

$\alpha_n^{max}$, $\alpha_n^{max}$, $\beta_n$, $\beta_p$, $E_n^{crit}$ and $E_p^{crit}$ are material dependent. The two critical fields are typically in the range of $10^7$ and $6\times10^8$ $\mathrm{Vm}^{-1}$ and determine the field strength at which this mechanism becomes important. Simulation of devices with this model included shows that it generally has little effect on the current density because the electric field strengths are not large enough. Also, because of the appearance of carrier currents in the expression for this type of generation, it can lead to instability in the numerical simulation of devices. Most of the time it can be ignored, but for devices under very large bias (or more accurately electric field) it should be included. If it is included into the simulation by the user, the program calculates the system equations without this mechanism included until the system has come close to convergence. If the system appears to be converging then this mechanism is then included automatically by the program.

## 6.8 Discretisation of System Equations

The three transport equations (Poisson's equation and the two current continuity equations) are discretised by means of taking finite differences. When only considering

steady-state conditions with a DC applied bias the three continuity equations in discrete form become:

$$\frac{D(M)-D(M-1)}{h'(N)} = q\left(N_D^+(N)-n(N)+p(N)-N_A^-(N)\right) \quad (6.67)$$

$$\frac{J_n(M)-J_n(M-1)}{h'(N)} = -(R(N)-G(N)) \quad (6.68)$$

$$\frac{J_p(M)-J_p(M-1)}{h'(N)} = (R(N)-G(N)) \quad (6.69)$$

It has long been recognised that stability issues arise when inverting the Jacobian Matrix needed to update system variables at each iteration. Kurata [67] gives a very informative discussion of this. He shows that instability occurs due to the Jacobian matrix not being strictly diagonally dominant when straight forward finite difference equations are used for the current density equations.

The condition of a matrix being strictly diagonally dominant is required to ensure that it is nonsingular and can therefore be inverted. Kurata [67] shows that if the potential variation between two points is too large then singularity of the Jacobian is likely to occur and numerical instability is introduced into the system. One method of addressing this problem is to decrease the spacing between points in order to keep the potential variation small. This method means that the number of nodes must increase and therefore the number of system variables and equations also increases which can be prohibitive. In addition it does not address the fundamental issues at hand.

Scharfetter and Gummel [78], in their ground-breaking paper on the subject, solved this problem by taking an integral form of the current continuity equations. By keeping the electric field, mobility and current density constant and integrating the continuity equations between two main points the integral form of the current equations is obtained. In this form the deviation from a strictly diagonally dominant Jacobian is negligible. This ensures nonsingularity and convergence of the equations.

The Scharfetter-Gummel form of current continuity equations for a $pn$ diode is given

by [67, 78, 77]:

$$J_n(M) = \; - \; \frac{qD_n}{h(M)} \left\{ n(N+1)B(\lambda) - n(N)B(-\lambda) \right\} \tag{6.70}$$

$$J_p(M) = \; \frac{qD_p}{h(M)} \left\{ p(N+1)B(\lambda) - p(N)B(-\lambda) \right\} \tag{6.71}$$

where $B(\lambda)$ is the Bernoulli function:

$$B(\lambda) = \frac{\lambda}{\exp(\lambda) - 1} \tag{6.72}$$

and the argument of the Bernoulli function, $\lambda$, is given by:

$$\lambda = \frac{q\left(\psi(N+1) - \psi(N)\right)}{k_B T} \tag{6.73}$$

These equations are valid for a simple device such as an isothermal, non-degenerate, homogeneous $pn$ diode, only. They do not handle more complex devices such as degenerate, non-isothermal heterostructures. In order to model these devices, the transport equations given in section 6.5.1 must be transformed into a similar form.

By following the treatment given by Smith and Rohatgi [74] the Scharfetter-Gummel form of carrier densities due to drift-diffusion given by Eq. 6.51 is given by:

$$J_n = \frac{\mu \Lambda_c k_B T_e}{h(M)} \left[ n(N)B(-\lambda_n) - n(N+1)B(\lambda_n) \right] \tag{6.74}$$

where

$$\lambda_n = -\frac{\Delta E_c}{\Lambda_c k_B T_e} - \left[ \left( \frac{5}{2} + S \right) \frac{\Omega_c}{\Lambda_c} - \frac{3}{2} \right] \frac{\Delta T_e}{T_e} + \frac{3}{2} \frac{\Delta m_e}{m_e} \tag{6.75}$$

A similar expression is available for holes.

## 6.9 Simulating Devices Under an Applied Bias

When a first-generation device (with structure as given in Fig. 6.4) is put under an applied bias the three coupled continuity equations given in section 6.7.1 need to be solved simultaneously at each node. Now there are 3 equations and 3 unknown variables at each node.

In calculating the Jacobian Matrix there are $\{3\times(N-2)\}\times 3\{\times(N-2)\}$ entries to be calculated (the partial derivative of each equation with respect to each variable). Similarly to solving Poisson's equation by itself, most of these entries will be zero. Only 9 derivatives

per equation need to be calculated. In order to optimise the code for speed, the equations and variables can be arranged so that the Jacobian matrix is *Band Diagonal*. That is, only the diagonal and 4 entries to the left and right of the diagonal are non-zero. This greatly reduces the time needed to calculate the Jacobian. By using special algorithms for the inversion of band-diagonal matrices [46] the time taken for the inversion of the Jacobian matrix can also dramatically reduced. Because of the absence of redundant zero entries the accuracy of the inversion due to round-off or truncation error is also increased.

### 6.9.1 Band Diagram of a First-Generation Device under an applied Bias

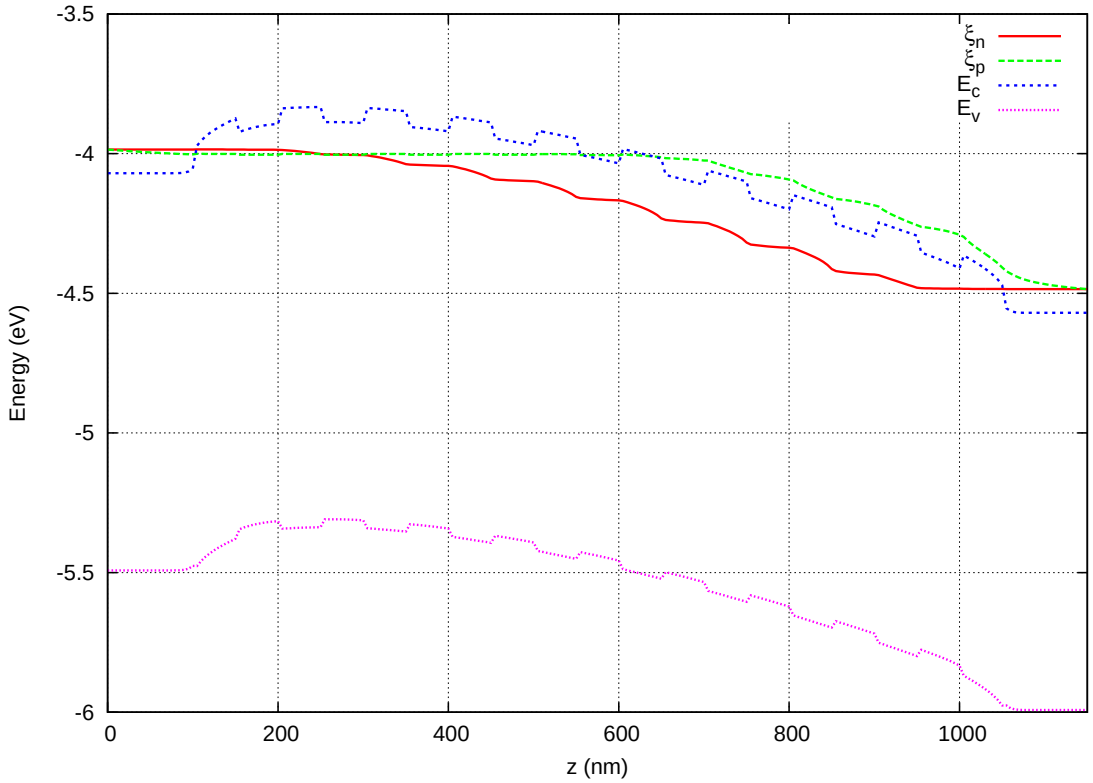


Figure 6.8: Band Diagram of 10-barrier heterostructure under applied bias of 0.5 V using Fermi-Dirac Statistics (active region only shown)

By applying an arbitrary applied bias, $V_{app}$ to the right-most node of the device, so that:

$$\psi_N = \psi_0 + V_{bi} + \frac{qV_{app}}{k_B T_N^{Lat}} \tag{6.76}$$

the three coupled continuity equations at each node are solved simultaneously by using

a multi-dimensional Newton-Raphson scheme. The band-diagram of the device with the structure shown in Fig. 6.4 under an applied bias of 0.5 V is shown in Fig. 6.8. Only drift-diffusion current is considered here so as to enable a comparison to when thermionic emission is included. Now the quasi-Fermi levels ($\xi_n$ and $\xi_p$) are no longer constant and no longer equal except at the two end nodes where the boundary conditions are set. The quasi-Fermi levels at the right-most node are equal and 0.5 V below the levels at the left-most node.

### 6.9.2 Effect of Thermionic Emission on the Band-Structure of a First-Generation Device under an applied Bias

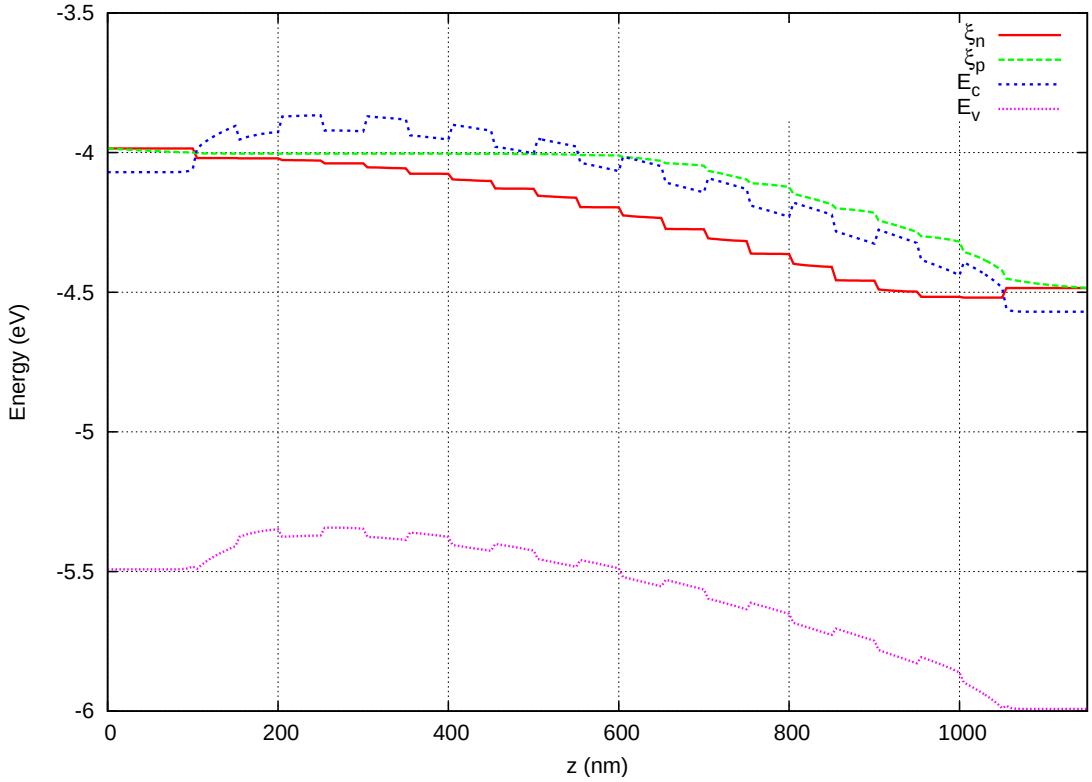


Figure 6.9: Band Diagram of 10-barrier heterostructure under applied bias of 0.5 V using Fermi-Dirac Statistics (active region only shown)

Fig. 6.8 shows the band-diagram of a first-generation device when only drift-diffusion current is considered. By including thermionic emission into the model it is seen in

Fig. 6.9 that it considerably alters the bands structure, and therefore operation, of the device. By comparing Fig. 6.8 and Fig. 6.9 it can be seen that thermionic emission does not change the conduction or valence bands of the device significantly. On the other hand there is quite a substantial change in the quasi-Fermi levels of the carriers. The inclusion of thermionic emission causes the quasi-Fermi level of the majority carriers (electrons) to change abruptly at the heterojunction interfaces. The quasi-Fermi level change of the minority carriers (holes) is also abrupt, but to a lesser extent. This means that thermionic emission is less important to minority carriers. This is expected because the barrier height (the energy difference between hole quasi-Fermi level and valence band) is much larger for minority carriers and fewer will have sufficient energy to overcome the barrier and contribute to thermionic current.

### 6.9.3 Recombination Rates of a First-Generation Device under an applied Bias

Fig. 6.10 shows the two dominant generation-recombination rates for a first-generation device. The top plot shows rates when only drift-diffusion is considered. The bottom plot includes thermionic emission.

Shockley-Read-Hall recombination is 2 orders of magnitude larger than optical generation-recombination. Auger recombination is 3 orders of magnitude smaller again and so is not shown on the plot. It is expected that this type of recombination should be small because it is only significant when there are large populations of both holes and electrons. Because GaAs is a direct band-gap material optical generation is significant.

When a bias of 0.5 V is applied to the device, the maximum electric field in the device is around $6\times10^6$ V/m. For GaAs, the critical field values for impact ionisation are around $7\times10^7$ V/m. This gives a large negative exponential value in Eq. 6.66 resulting in a very small (approximately zero) value for impact ionisation, so this type of generation is unimportant.

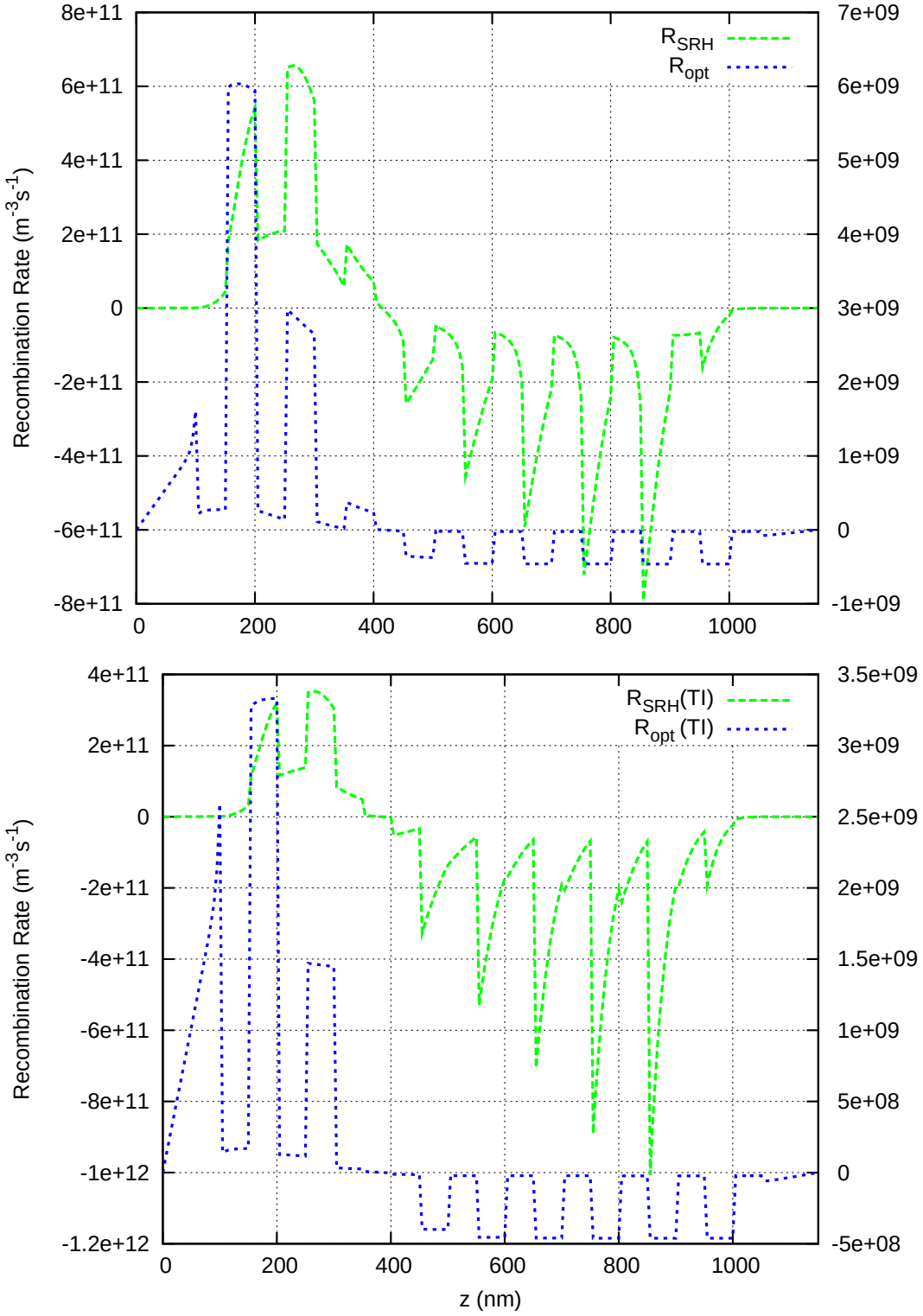


Figure 6.10: Recombination rates of first-generation under applied bias of 0.5 V. The top plot only includes drift-diffusion, the bottom plot also includes thermionic emission. $R_{SRH}$ uses left axis, $R_{opt}$ uses right axis.

*TISim* also calculates the current density of both majority and minority carriers through the device for a given bias. Due to non-zero recombination-generation rates, the currents are not necessarily continuous throughout the device (see Eq. 6.56 and Eq. 6.57). At an applied bias of 0.5 V and with thermionic emission included the electron current density is around $-4\times10^{6}$ Am$^{-2}$. This corresponds to an electron flux of around $2.5\times10^{25}$ m$^{-3}$s$^{-1}$. The net generation-recombination rate is of the order of $10^{12}$ m$^{-3}$s$^{-1}$ so the effect is negligible for electron current density and it can be assumed constant throughout the device. The minority current density, on the other hand, is much smaller (around $-10^{-14}$ Am$^{-2}$) and corresponds to a flux of around $10^{6}$ m$^{-3}$s$^{-1}$. Thus the effect of recombination-generation on hole transport becomes important and the hole quasi-Fermi level changes. But, because the hole current density is around 20 orders of magnitude smaller than electron current density, this change is negligible to the total current density.

These results show that for the first-generation devices studied, recombination-generation rates do not significantly effect their overall operation, either electrically or thermally, because they only significantly effect minority carriers which do not contribute significantly to device operation. They can therefore be safely ignored.

## 6.10 $I$-$V$ characteristics of a First-Generation Device

With the inclusion of current continuity equations into the model the electrical current densities of the device can be calculated for different applied biases. This will allow comparison between the model and experimentally obtained $I$-$V$ characteristics.

Chapter 7 will use the generated $I$-$V$ curves to analyse the experimental results of the three generations of devices produced so far.

## 6.11 Conclusion

This chapter has discussed the analysis of experimental $I$-$V$ characteristics using the previously derived model for multi-barrier devices. It was found immediately that the simple model developed in previous chapters neglected some very important physics in the devices thus inadequately modelling the experimental results correctly.

In order to address this, the first step was to include barrier height lowering due to space charge accumulation. This was initially done by using image force lowering as an approximation of the space charge potential near heterojunction interfaces. Section 6.1 showed that this approximation did not accurately describe electron transport in the devices. By solving Poisson's equation it was shown in Section 6.3.1 that this was most likely caused by non-linear variation of the space-charge potential in the device. In order to use image-force lowering as an approximation to the space-charge potential it is required that the space-charge potential varies linearly near the heterojunction interfaces.

To further investigate the electrical transport in the devices a more robust model of the devices using finite differences was developed. By doing so a better approximation for the energy band diagram and space charge accumulation was included. In addition to this current due to drift-diffusion away from the heterojunction interfaces was included, with current due to thermionic emission near the interfaces used instead. Recombination and generation rates were also included for completeness but were found to have little effect on the current densities in the devices studied.

The next chapter will compare the new device model to the experimentally obtained electrical characteristics and it will be seen that there is now much better agreement. Also, because of the extra complexity of the model, much more information about the devices is available.

# CHAPTER 7

# ANALYSIS OF EXPERIMENTAL RESULTS AND DESIGN OF THIRD-GENERATION DEVICES

## 7.1 Simulation of an 'Ideal' First- and Second- Generation Device

The first- and second- generation devices are nominally identical except that the second-generation devices also have metallisation deposited directly onto the substrate. The purpose of this was to experimentally determine the effect that the substrate, metallisation and contacts have on the device. Having said this, the $I$-$V$ characteristics of the first two generations of devices should be very similar.

### 7.1.1 Effect of Substrate on Device Operation

Fig. 7.1 shows the input file for an 'ideal' first- or second- generation device with the 450 $\mu$m $n^+$-GaAs substrate included.

Fig. 7.2 shows the band diagram of this device with an applied bias of 1 V. This is done to see the effect that the substrate has on the device because the active region only constitutes a small fraction of the total width of the device. The top plot in Fig. 7.2 illustrates this point. The active region is shown as a vertical line at the right side of the top plot. The bottom plot shows only the active region of the device. The total width of the device is $\sim$451 $\mu$m while the active region is only $\sim$1 $\mu$m.

Fortunately, because the substrate is highly doped ($\sim 2\times10^{18}$ cm$^{-3}$) most of the bias is dropped over the active region. This means that negligible energy is dissipated in the substrate and so it has little effect on the overall operation of the device. When the current density at an applied bias of 1 V is calculated with and without the substrate

```
Mat=AlGaAs x=0.0  Nd=2e24 Width=450000 Step=1000
Mat=AlGaAs x=0.0  Nd=2e24 Width=100    Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.07 Nd=0    Width=50     Step=5
Mat=AlGaAs x=0.0  Nd=2e24 Width=100    Step=5
AppliedBias=0
```

Figure 7.1: Listing of ‘Ideal’ First- and Second- Generation Device with 450 $\mu$m Substrate Included

present it is found that the presence of the substrate only lowers the current density by around 3.8%, even though it makes up over 99.7% of the total device width. This is the expected behaviour as the substrate has almost metallic properties due to its high doping level.

### 7.1.2 Simulated $I$-$V$ Characteristics

Fig. 7.3 shows the simulated $I$-$V$ characteristics of an ‘ideal’ first- and second- generation device. For completeness all available generation-recombination rates (Optical, SHR, Auger and Impact Ionisation) are included although they have little effect on the generated curve, as discussed in section 6.9.3. Thermionic emission is included. This acts to limit the current density through the device.

The generated $I$-$V$ characteristics show that the presence of the substrate has little effect on the electrical operation of the device. The current density magnitude is still very

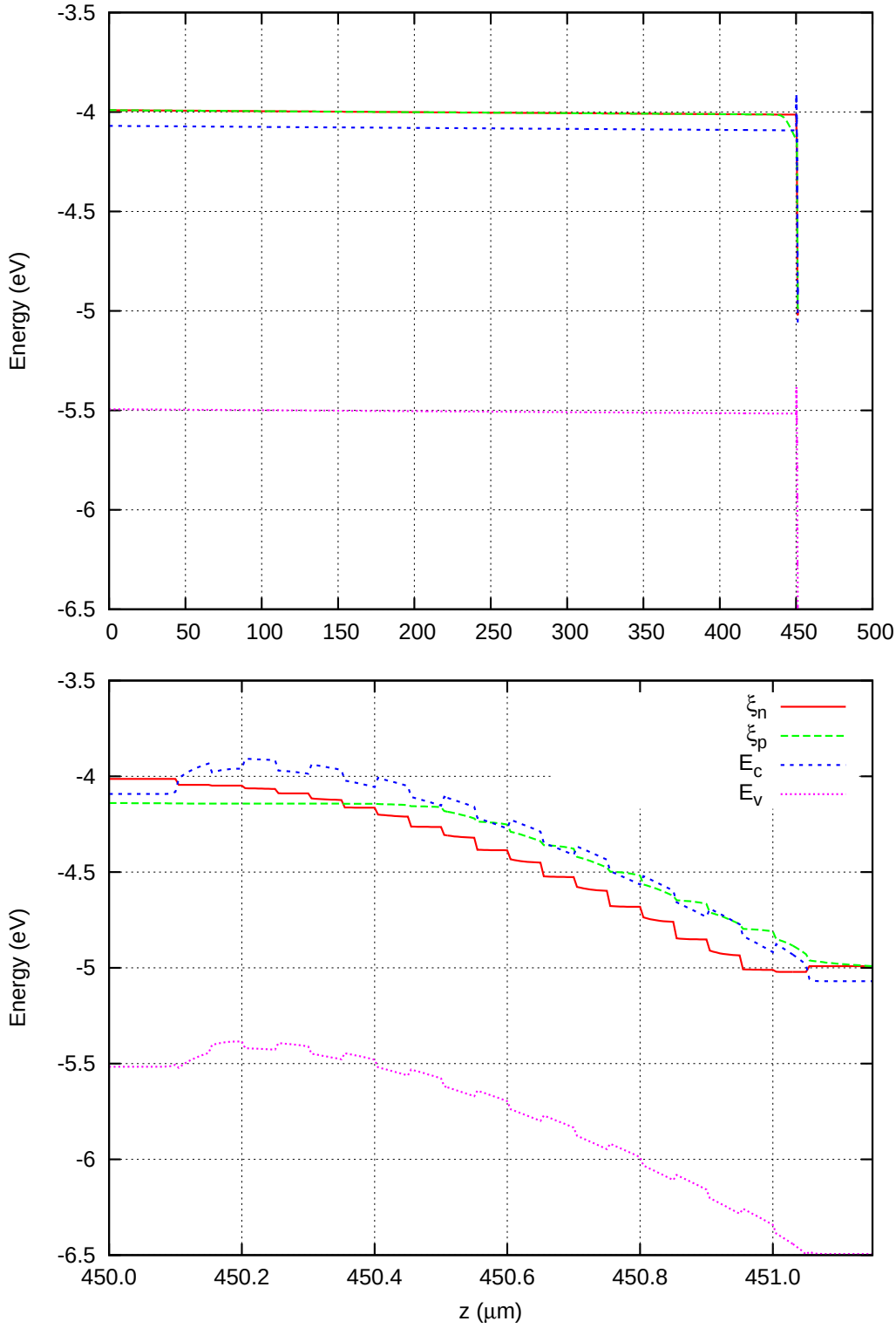


Figure 7.2: Energy-Band diagram of 'Ideal' Device. The bottom plot shows an expanded view of the active region of the device

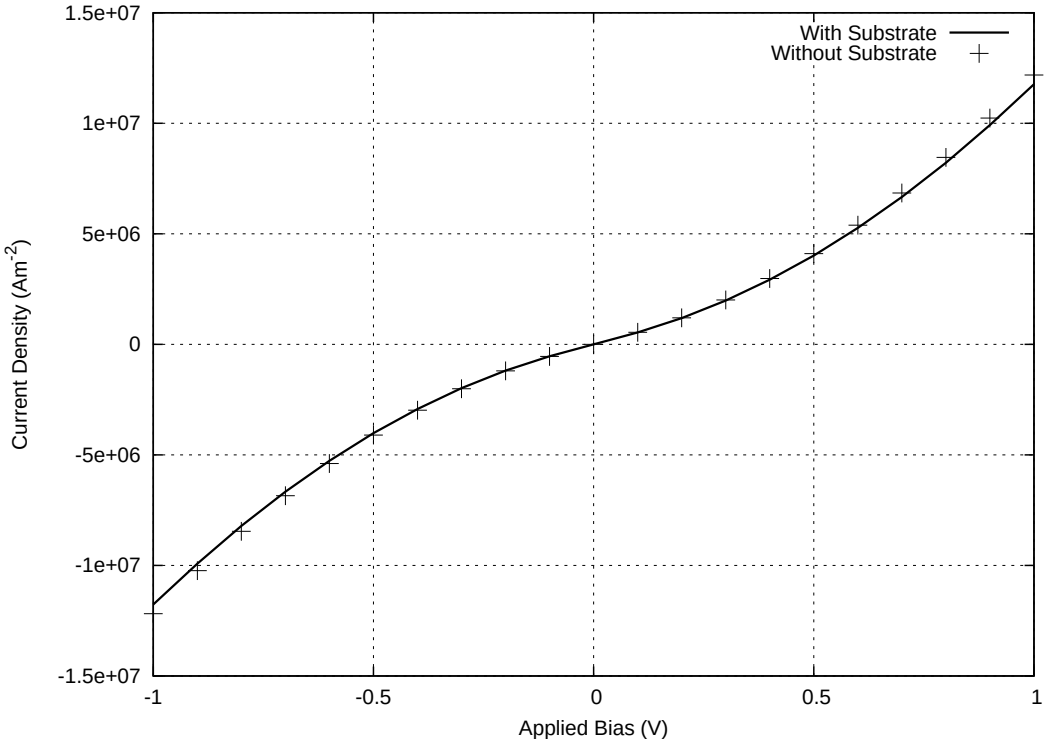


Figure 7.3: Electrical Current Density versus Applied Voltage for 'Ideal' first-generation device

large ($\sim 1.2 \times 10^7 \mathrm{Am}^{-2}$). For a device with diameter of 3 mm this equates to a current of $\sim$83 A. This is clearly much too large. The difference in magnitude could be attributed to many non-ideal effects such as dislocations in the lattice, oxidation of the surfaces or perhaps damage to the devices.

The next section will compare the experimental and simulation results.

## 7.2 Analysis of First- and Second- Generation Devices

Due to the fact that the real devices will not have a structure identical to the specifications given, the device structure used in simulation can be modified slightly until the correct trends become apparent in the simulated characteristics, when compared to the experimental results. If this is done correctly, then other characteristics of the devices can be determined theoretically, such as a better approximation for the energy band diagram of the device.

Negligible cooling was observed in the devices produced in this work. The reason that no cooling was observed in the first- and second- generation devices was because of a flaw in the design of these devices. The initial design was based on the assumption that the barrier height in the devices was simply the conduction band offset at the heterojunction interfaces. These devices were designed very early in the course of the work and the physics of the devices was not completely understood at that time. With derivation of Richardson's equation (Section 2.2) it is immediately obvious that the barrier height is actually measured between the chemical potential in the first material and the conduction band in the second material. Because of the use of undoped semiconductor in the active region of the device the chemical potential in this region was far in the band gap. Although the conduction band offset was designed to be around 70 meV (the barrier height required to achieve appreciable cooling at room temperature), the actual barrier height (measured between chemical potential and conduction band) was much larger[1]. This will be discussed further in the next sections.

Although no cooling was observed (or expected with this design flaw taken into consideration), thermionic emission still considerably influences the electrical characteristics of the devices. The experimental $I$-$V$ characteristics still allow for a great deal of information to be learnt about the devices. They allow for a test to determine whether the models used described the devices sufficiently.

The previous chapter showed that indeed the simple model used initially was inadequate for describing the electrical characteristics of the devices. Even when barrier height lowering due to the space charge potential was included in the device the model did not sufficiently describe the devices. A more detailed model was developed and the next sections will compare these results with those obtained experimentally.

[1] In modelling these devices it is assumed that chemical potentials at different points can be set independently of each other and are dependent only on material parameters. They are then lined up by adjusting the built-in potential at each point. This causes the conduction band of the active region to be far away from the chemical potential of the doped regions, giving a large barrier height.

### 7.2.1 $I$-$V$ Characteristics

The experimental results in this section are taken from the Master's Thesis of Sueping Lee [49] who carried out the majority of experimental work on these devices. A great deal of electrical data was collected for these devices. Some of the results which should be in very close agreement with each other (because they had identical design) were different. Because InGa eutectic was used to physically contact the bottom of the device to the brass heatsink (see Fig. 5.3), and because of the possibility of oxidation, physical and therefore electrical contact may not have been sufficient in every case resulting in different results. The results presented and analysed here are those which were repeatable and similar between different devices and between the first two generations.

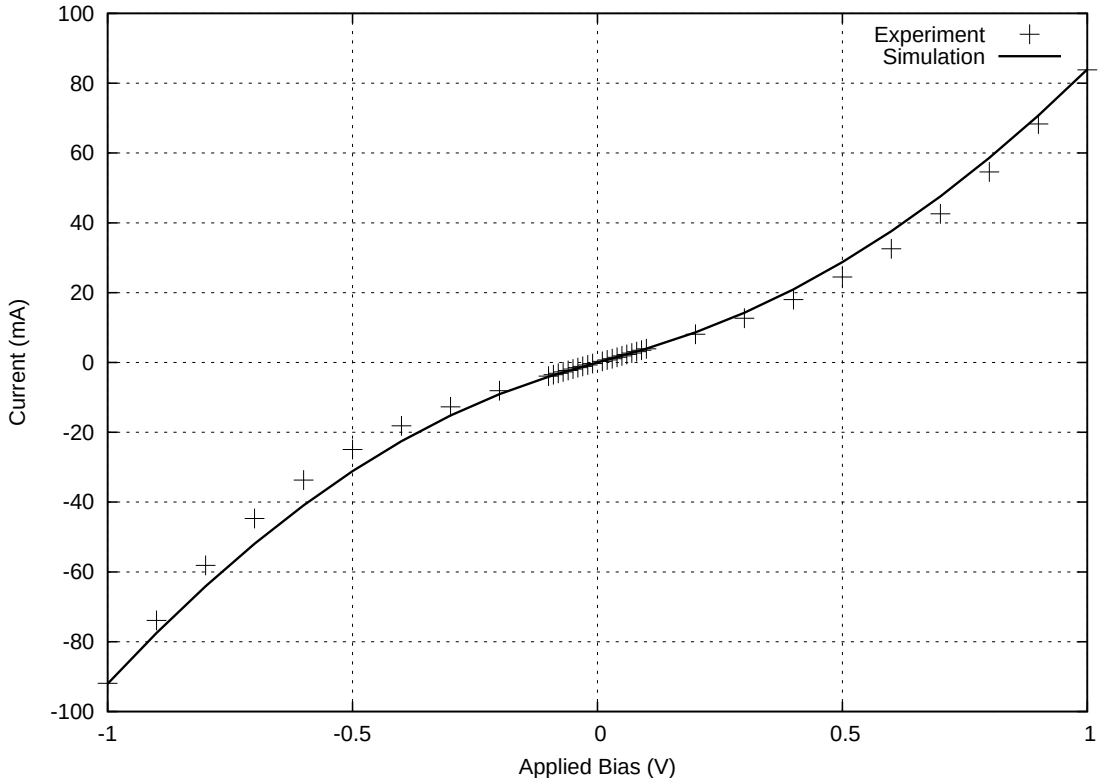


Figure 7.4: Calculated and Experimental current of first-generation device.

Fig. 7.4 shows the experimental results previously presented and results from the simulation. The experimental results shown are the same as those in Fig. 6.2. The device

structure used is similar to that shown in Fig. 7.1 but with the Aluminium concentration of the second electrode changed from the nominal 0% to 3%. All other aspects of the structure are unchanged.

This change of Al concentration to 3% is used to account for the non-symmetric nature of the experimental curve. The current is greater for a negative applied bias than for the corresponding positive bias. As previously stated the actual device structure will be different to the 'ideal' device design. This 3% concentration could be due to diffusion of Al during fabrication. A better approximation may be to have small regions where the Al concentration is graded between heterojunction interfaces, but this approximation models the experimental results sufficiently.

With a better approximation for the 'real' device structure obtained by comparison with experimental $I$-$V$ characteristics an analysis of the energy band diagram is now possible.

### 7.2.2 Band Diagram of 'real' Device

With the device structure used to generate the $I$-$V$ characteristics in the last section we are at last able to obtain an approximation of the energy band diagram of a 'real' first- or second- generation device. This is shown in Fig. 7.5. The region of 3% Al concentration is shown on the plot, along with the actual barrier height that is necessary for electrons to overcome to contribute to thermionic current.

The design flaw previously discussed becomes immediately obvious. Rather than the barrier height being around 70 meV as desired it is closer to 215 meV at its largest value. As mentioned this is because the barrier height was initially assumed to be simply the conduction-band offset at the heterojunction interface rather than the difference between chemical potential and conduction band energy.

Using the model developed in chapter 4 to determine the maximum cooling possible for a 10-barrier device made from GaAs/AlGaAs with barrier height of 215 meV, the maximum cooling for this device is expected to be 0.00236493 K, or 0.00751894 K with

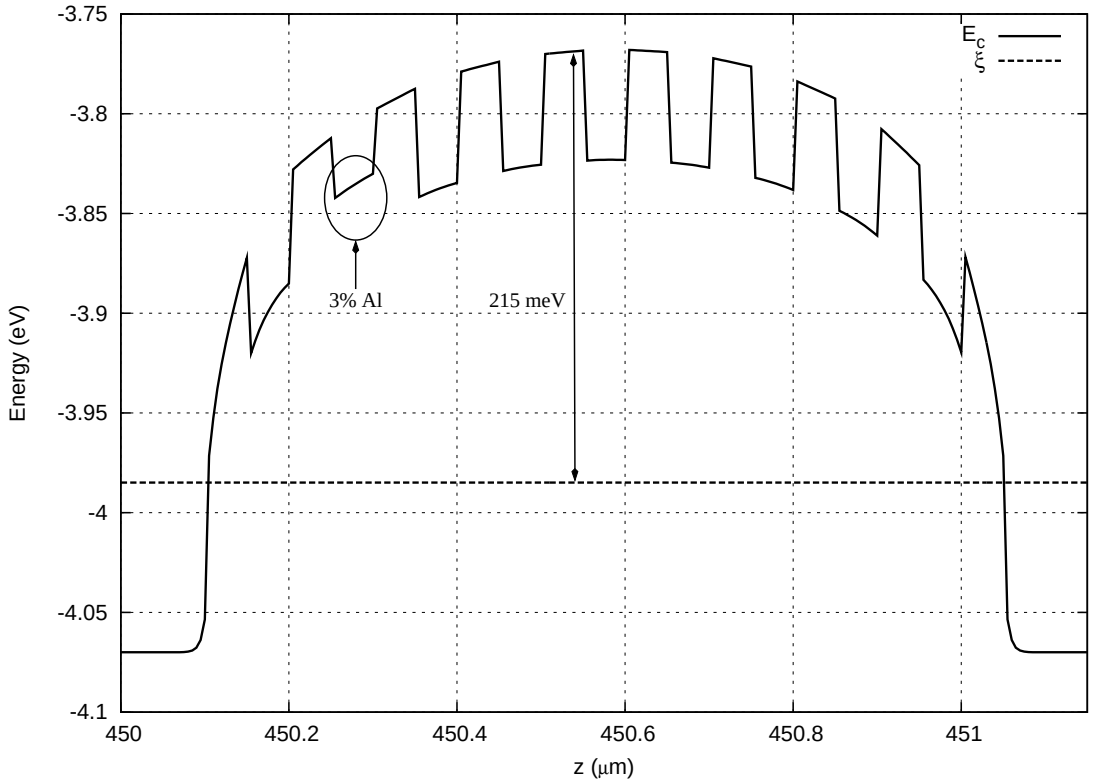


Figure 7.5: Energy Band Diagram of ‘real’ first- or second- generation device.

phonon scattering included. Thus, little or no observable cooling is expected from these devices.

### 7.2.3 Conclusion

With the development of the more robust model in chapter 6 better agreement between experimental and theoretical results was achieved. The device structure of an ‘ideal’ first- or second- generation device was modified slightly to account for the non-symmetric nature of the experimental $I$-$V$ characteristics.

With this modified device structure more detailed information about the first two generations of device was available from the theoretical model. It was found that there are regions of non-linear space-charge accumulation in the device which probably accounts for the inability of the simple model developed in chapter 4 to successfully describe the electrical characteristics of these devices. The new model showed that there is substantial

space-charge near the heterojunction interfaces which modifies the energy-band diagram of the device. The simple model derived earlier initially assumed barrier heights that are independent of applied bias. A slight modification of this model tried to model barrier heights dependent on applied bias (or more accurately electric field) by approximating the space-charge potential by image-force lowering. This approximation assumed that the space-charge varies linearly near the heterojunction interfaces and so was unsuccessful at modelling the characteristics sufficiently.

The model used to analyse the experimental results in this chapter showed that, although it accounts for over 99% of the total width of the devices, the substrate had little effect on the overall electrical operation of the devices because of its metallic-like behaviour due to high donor concentration.

Perhaps the most important result obtained from the simulation results was the obvious design flaw of the first two generations of device. That is, that the actual barrier heights were closer to 215 meV rather than the 70 meV required for appreciable cooling at room temperature.

With these results and the new model available it was possible to design new devices that would address this flaw. The large barrier heights were caused by using undoped material for the devices. This caused the chemical potential in the active region of the devices to be far in the band-gap, away from the conduction band energy at the heterojunction interfaces.

The third generation of devices aimed to address this problem by using high levels of doping in the active region, effectively pushing the chemical potential near or above the conduction band. Modelling the band structure with the new model shows that these devices have barrier heights closer to the required value. The remainder of this chapter will deal with the design and analysis of these new devices.

```
Mat=AlGaAs x=0.0  Nd=1.5e24 Width=300000 Step=10000
Mat=AlGaAs x=0.0  Nd=1.5e24 Width=500   Step=25
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.07 Nd=5e21   Width=90    Step=5
Mat=AlGaAs x=0.0  Nd=5e22   Width=100   Step=5
Mat=AlGaAs x=0.0  Nd=1.5e24  Width=400  Step=20
AppliedBias=0
```

Figure 7.6: Listing of ‘Ideal’ Third-Generation Device with 450 $\mu$m Substrate Included

## 7.3 Design of Third-Generation Devices

The third-generation devices are grown on an MTI Si-Doped GaAs Single Crystal Substrate using Molecular Beam Epitaxy (MBE) as opposed to MOCVD for the first two generations. According to the manufacturer’s specifications the carrier concentration of the substrate is (5∼15) $\times$ $10^{17}$ cm$^{-3}$. The device file in Fig. 7.6 assumes that the substrate carrier concentration is 15 $\times$ $10^{17}$ cm$^{-3}$. A lower value will effect the electrical characteristics only be reducing the magnitude of the current density. The substrate is now only 300 $\mu$m as opposed to 450 $\mu$m for the first two generations.

The third-generation device design is similar to the first- and second- generation design. Notable differences are that the electrodes and barriers are now doped with 5 $\times$ $10^{17}$ cm$^{-3}$ and 5 $\times$ $10^{16}$ cm$^{-3}$ donors, respectively. This acts to move the chemical potential closer to the conduction band thus reducing the barrier height of the majority carriers, electrons.

The barrier widths have been increased from 50 nm in the first-generation design to 90 nm in this design. This value is close to the value of the mean-free path of electrons in GaAs calculated in section 3.5.1. Although the barrier material is now doped AlGaAs rather than pure GaAs it is assumed that the mean-free path is similar to this previously calculated value. By increasing the barrier width we reduce the thermal conductance and therefore conduction heat flow in the devices.

### 7.3.1 Band Diagram of 'Ideal' Third-Generation Device

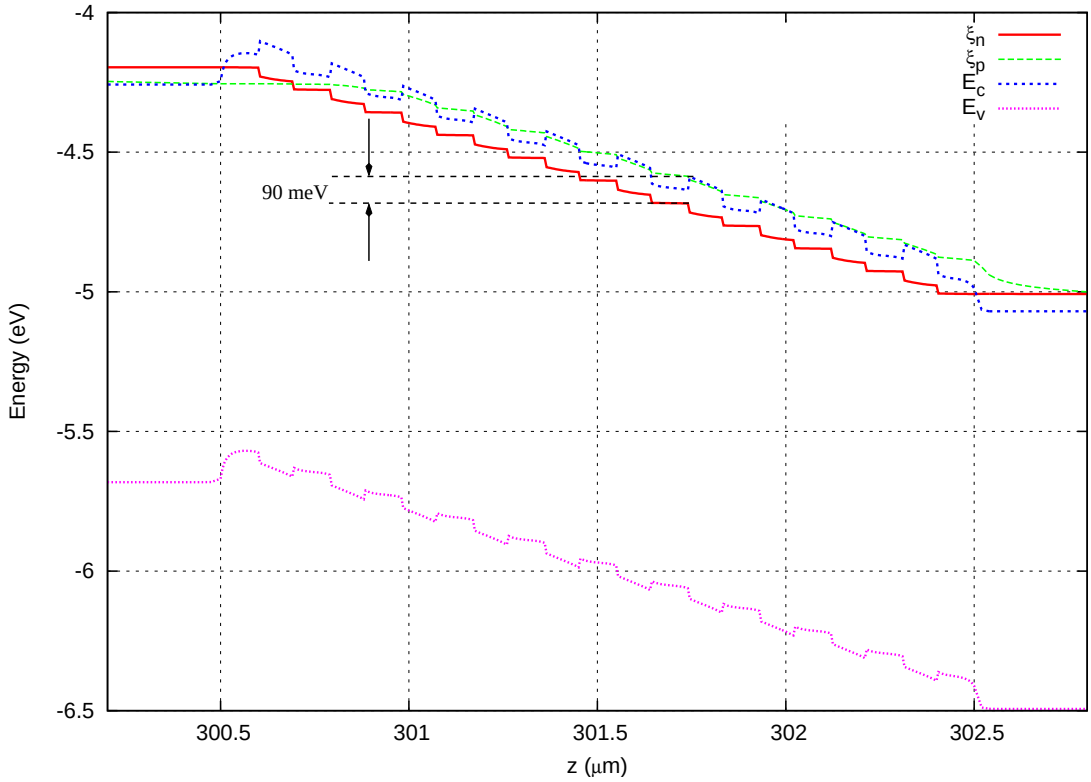


Figure 7.7: Band diagram of 'Ideal' third-generation device under 1 V applied bias.

Fig. 7.7 shows the band diagram of an 'Ideal' third-generation device under an applied bias of 1 V. The 300 $\mu$m substrate is not shown. Because the active region now contains doped material proportionally more energy is dissipated in the substrate. In the first two

generations the active region had a much larger resistivity because of the absence of carriers. The electron quasi-Fermi level drops from -4.01 eV to -4.20 eV across the substrate meaning that around 20% of the energy input to the system is dissipated in the substrate. This will mean that any cooling in the active region may be masked by substantial joule heating in the substrate. This joule heating could be reduced by substrate thinning but the facilities were not available for these devices. One positive aspect is that the substrate width of the third-generation devices is only 67% of the previous generations.

Inspecting the electron quasi-Fermi level and conduction band in Fig. 7.7 it is found that the barrier heights for electrons at each heterojunction interface are similar and equal to around 90 meV. This is a much better result than for the previous device designs.

### 7.3.2 Theoretical I-V Characteristics of an 'Ideal' Third-Generation Device

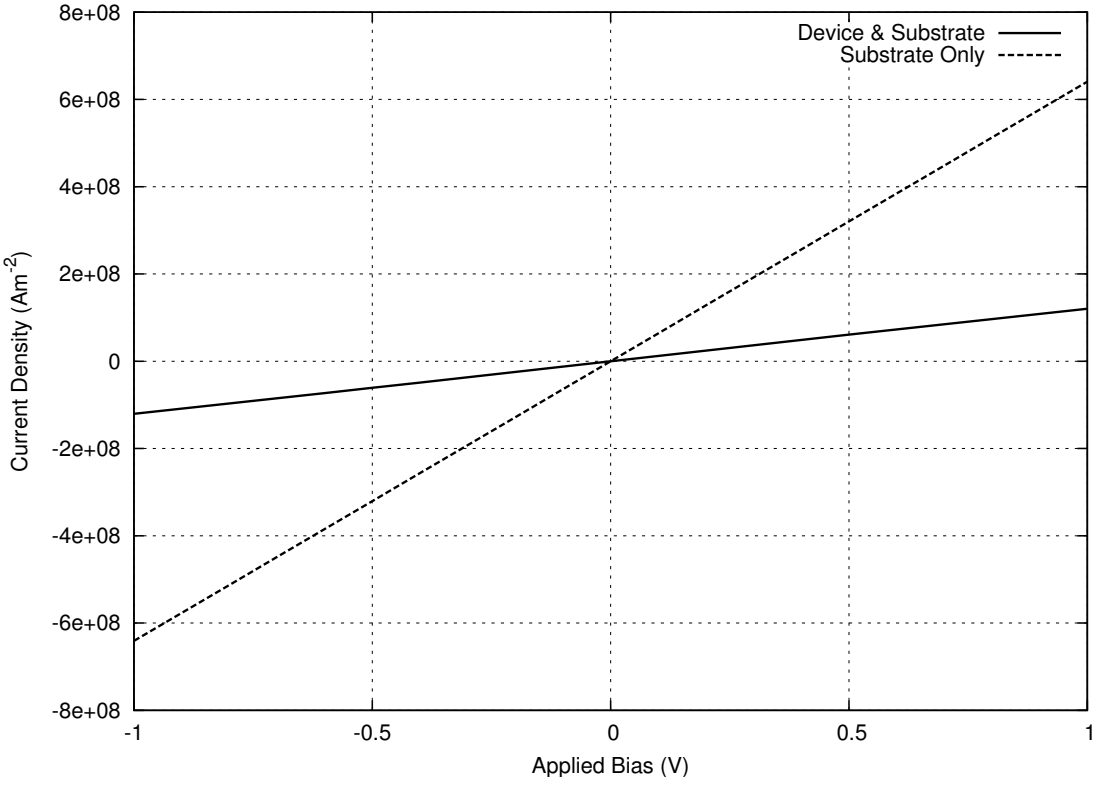


Figure 7.8: Simulated I-V characteristics of 'Ideal' third-generation device and substrate alone.

The theoretical $I$-$V$ characteristics for an 'ideal' third-generation device are shown in Fig. 7.8. Also shown on the plot is the $I$-$V$ curve for the substrate alone. The magnitude of the current density due to the substrate alone is around 5 times larger than for the device and substrate together, agreeing with the statement in the last section that around 20% of the applied bias is dropped over the substrate.

## 7.4 Analysis of Third-Generation Devices

Comparison of the theoretical $I$-$V$ characteristics in Fig. 7.8 with the experimental results shown in Fig. 5.6 and Fig. 5.7 shows that there is no correlation between the two. The theoretical results show that the device has ohmic characteristics whereas the experimental results are highly nonlinear and also nonsymmetric. This asymmetry may be due to the substrate and contacts of the real device. The experimental results for the substrate alone support this supposition. These results shown in Fig. 5.8 can now be used to analyse the devices themselves.

Comparing the magnitude of the current at an applied bias of $-700$ mV for the substrate only ($\sim -19.73$ mA) to the current for the device ($\sim -4.74$ mA) shows that the substrate has an impedance of around 20% that of the device and substrate together. This at least agrees with the theoretical results presented in section 7.3.2. For positive applied bias this does not hold.

### 7.4.1 Equivalent Circuit

It was assumed that by using the $I$-$V$ characteristics of the substrate alone and of the substrate and device together that the $I$-$V$ characteristics of the device alone could be found. Referring to Fig. 7.9 we assume that the current, $I$ flowing in both circuits is equal. The applied biases across both circuits ($V_a$ and $V_b$) will be different. Then, the current through circuit a) is equal to:

$$I = \frac{V_a(I)}{X_s(I)} \tag{7.1}$$

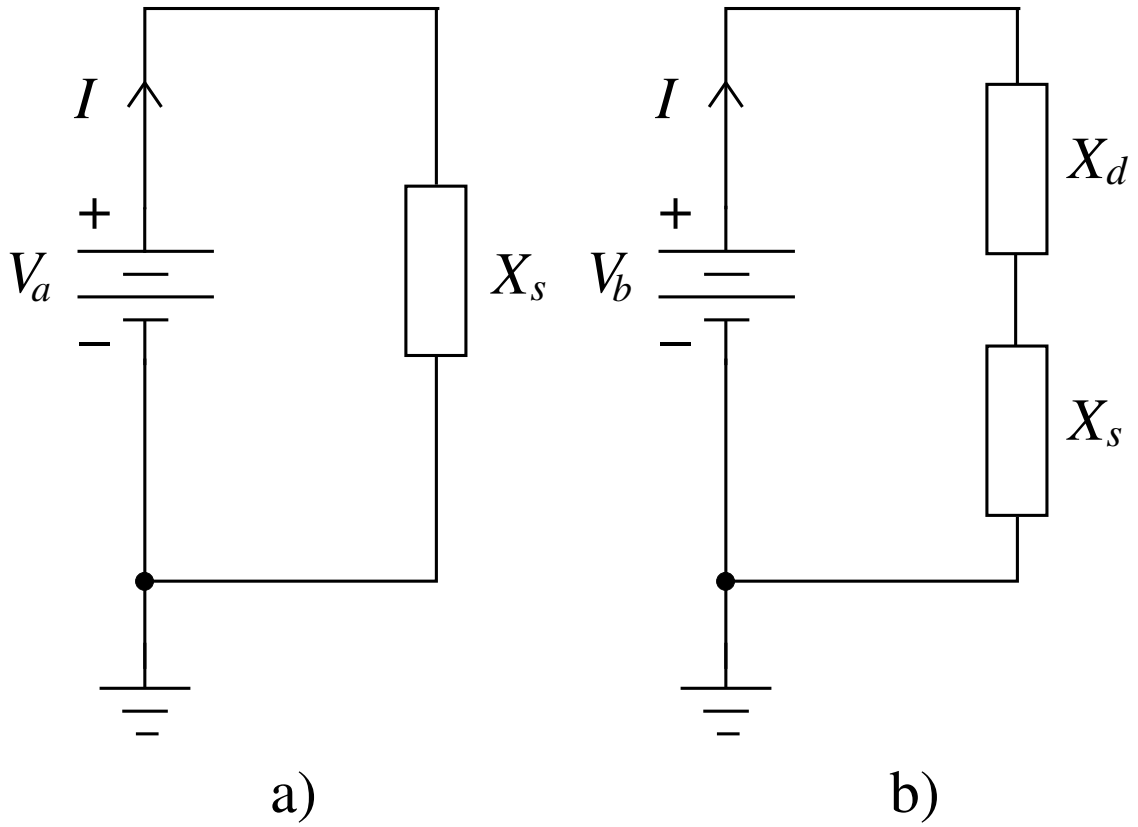


Figure 7.9: Equivalent Circuit of a) Substrate alone and b) Substrate and Device together.

and the current flowing in circuit b) is:

$$I = \frac{V_b(I)}{X_s(I) + X_d(I)} \tag{7.2}$$

Equating these two expressions gives the impedance of the device alone, $X_d(I)$, for a given current, $I$, which can be used to calculate the $I$-$V$ characteristics of the device alone:

$$V_d(I) = I \cdot X_d(I) = I \cdot \left[ \frac{V_b(I)}{V_a(I)} X_s(I) - X_s(I) \right] \tag{7.3}$$

Using the fact that

$$X_d(I) = \frac{V_a(I)}{I} \tag{7.4}$$

we get:

$$V_d(I) = V_b(I) - V_a(I) \tag{7.5}$$

Aside from this formal derivation it is easy to see from circuit b) in Fig. 7.9 that the voltage drop across the substrate alone for a given current is simply equal to $V_a(I)$ so the

voltage across the device alone is that given by Eq. 7.5.

In order to calculate $V_d(I)$ from Eq. 7.5 we need the voltage across the substrate alone for a current equal to the current given for the substrate and device together. Generally this is not available directly but is found by linear interpolation.

For this analysis it is more useful to plot $I$ vs. $V$ in order to verify that, for a given current, the interpolated voltage across the substrate added to the calculated voltage across the device alone is equal to the voltage measured across the substrate and device together.

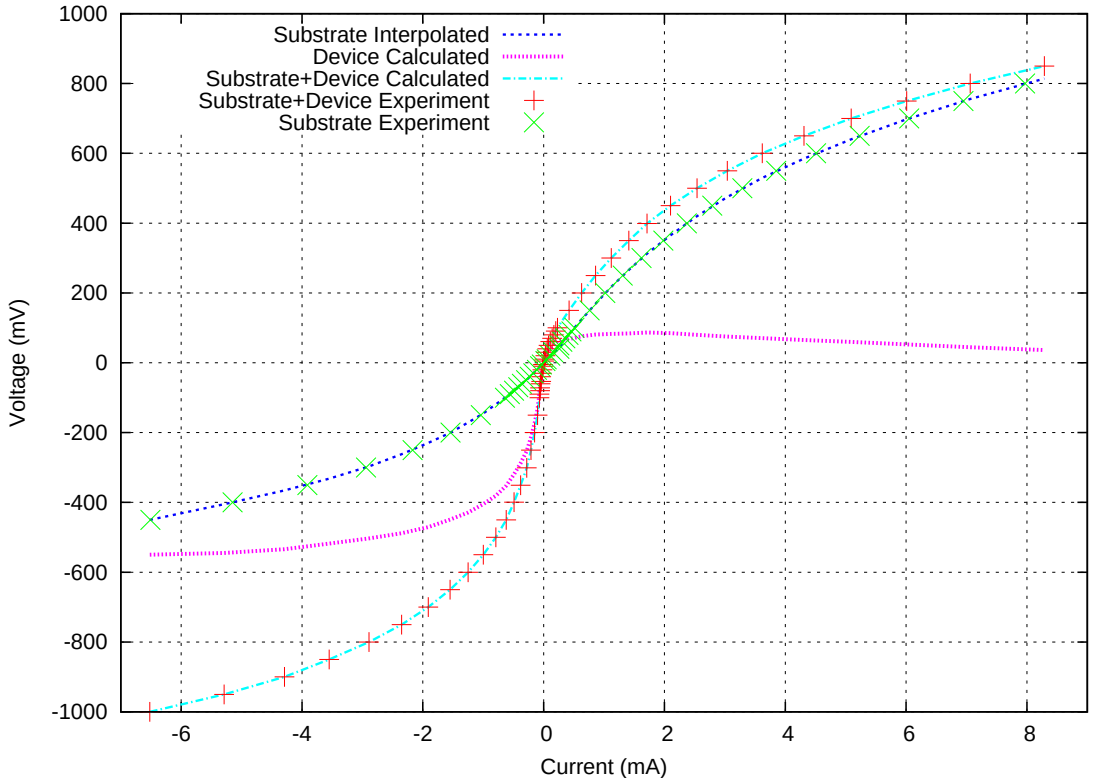


Figure 7.10: Calculated V-I characteristics of third-generation device alone.

Fig. 7.10 shows the calculated $V$-$I$ characteristics for a third-generation device alone. Also shown is the interpolated $V$-$I$ characteristics of the substrate alone used to calculate the device characteristics. As a comparison the experimental results for the substrate alone are shown and agree with the interpolated results. In addition to this the calculated

$V$-$I$ characteristics of the substrate alone added to those of the device alone are compared to the experimental $V$-$I$ characteristics and they also agree, confirming that the interpolation scheme is functioning correctly.

### 7.4.2 Analysis of Calculated $V$-$I$ Characteristics of Equivalent Circuit

The calculated $V$-$I$ characteristics of the device alone shown in Fig. 7.10 highlight that the assumptions of the equivalent circuit are incorrect. For negative currents the characteristics behave well, but for positive currents the calculated voltage across the device alone actually decreases with an increase in current. This indicates that the contacts on the substrate alone are different to the contacts grown onto the devices. Two possible reasons for this are that either the samples became damaged after fabrication or that the surfaces of the samples may have become contaminated or oxidised before the metal contacts were grown onto the samples. By comparing the experimental results of these devices with those of first-generation devices it is likely that oxidisation may be the reason.

### 7.4.3 Comparison Between Third- and First- Generation Devices

Referring to Fig. 5.4 and Fig. 5.6 the current for the 6 mm first-generation device at an applied bias of $-900$ mV is $-68.33$ mV. The corresponding current for the 3 mm third-generation device is $-8.7$ mV. Scaling the current for the 3 mm device (multiply by 4) gives a current of $-34.8$ mV for a corresponding 6 mm device. The scaled current between the two generations shows that the magnitude of the current for the third-generation devices is only half that of the first- and second- generation devices. If anything, the first-generation current should be significantly less than third-generation devices because of the use of undoped barriers and electrodes. By using undoped material there should be less carriers available so the overall current should be much less for first-generation devices. This means that some other mechanism must be inhibiting the flow of carriers in the third-generation devices.

The most likely cause of decreased current would be oxidation or contamination of sample surfaces before growth of metal contacts. The $I$-$V$ characteristics of the substrate alone show trends that are different to standard contacts such as Schottky barriers or ohmic contacts. Schottky barriers generally rectify the current whilst ohmic contacts give a linear variation of current with applied bias. The shape of the $I$-$V$ characteristics of the substrate alone shows that some non-standard contact is present in the samples. Oxidation or contamination of the sample surfaces would reduce overall current and likely change the shape of the $I$-$V$ characteristics so this seems the most likely cause. Damage to the samples after fabrication is unlikely due to the fact that the experimental results between different samples are very similar.

### 7.4.4 Conclusion

The $I$-$V$ characteristics of third-generation are not what are expected from theory. The shape of $I$-$V$ characteristics are different to what is predicted by the model. Because of the high doping through the device, the model predicts that the devices should have ohmic qualities. That is, they should vary linearly with applied bias. In addition to this, the magnitude of the current for a given bias is much less than it should be. Third-generation devices should induce much larger current than first-generation devices because of the higher doping carrier concentration throughout the device, meaning that more carriers should be available.

In order to examine the $I$-$V$ characteristics of the active region of the devices alone (without the substrate and contacts) an equivalent circuit was used to try and isolate these characteristics. As seen by section 7.4.1, this approach did not work. This signifies that the contacts directly on the substrate are not the same as those on the devices.

These results could mean that the surfaces may have been contaminated or become oxidised during fabrication. This would explain why the magnitude of the current is less than it should be and why the shape of the $I$-$V$ characteristics are different to theory. The

$I$-$V$ characteristics of the substrate alone do not appear to have the shape of common contact types such as ohmic or Schottky barriers. Schottky barriers will rectify the current in one or more directions of applied bias. The $I$-$V$ characteristics of the substrate alone do not have these trends. Contamination or oxidation of the samples could explain the results. One other possibility is that the devices themselves became damaged after fabrication, but this can probably be ruled out because the $I$-$V$ characteristics are very similar between samples. It is expected the characteristics between samples should be different if one or more of them had become damaged.

With a theoretically superior design for the third-generation devices it was hoped that cooling would be observed in these devices. Analysis of the electrical characteristics of these devices show that they do not behave as they should. This indicates that there may have been problems in the fabrication process. Taking the electrical results into account, it is understandable that no cooling was observed. Less than 0.5 K cooling was expected in these devices, and this was assuming that they were close to an 'ideal' device. Any fabrication problems would therefore significantly decrease the likelihood of cooling being observed. In order for cooling to be observed in the future the best solution is to use a better device design in which more cooling is expected. The next section will discuss a new device design that has theoretically higher cooling and, as a result, should increase the probability of observable cooling.

## 7.5 Towards a Better Device Design

### 7.5.1 Introduction

The results to this point have shown that GaAs-based systems will not perform well due to the high thermal conductivity of the material. New generations should be made of more suitable material. The devices in this work were made from GaAs primarily because of availability and also because GaAs-based thermionic devices had not been made before. In addition to using different material, a different device structure should also increase the

thermal resistivity of the device due to interface scattering and localisation of phonons.

### 7.5.2 Reducing Thermal Conductivity

As has been mentioned in previous sections of this thesis, a substantial amount of work has been carried out on measuring and modelling the reduced thermal conductivity of superlattices as compared to bulk properties. Huang et al. [21] showed that optical-phonon transport in semiconductor superlattices is significantly reduced if a Fibonacci sequence is used for the superlattice length rather than having structures with equal periods. The increased localisation of phonons may be due more to the simple introduction of random period lengths rather than the Fibonacci series specifically. By introducing random barrier widths into the system the phonon transport may be reduced because, treating the phonons as waves, standing waves will not be setup and propagate as easily. By doing so it is anticipated that by using multiple barriers, thermal conductivity will be reduced because of this and because of interface scattering.

### 7.5.3 Material Selection

From the results of chapter 4 it was found that n-type $Ga_{0.47}In_{0.53}As$ will make a much better material than GaAs, due to its relatively large effective mass, mobility and thermal resistivity. It is recommended that new devices should be made using this material for the barriers.

To ensure lattice constant matching and for easier fabrication the electrode material should also be made of $Ga_xIn_{1-x}As$. Because the band gap of this material increases with $x$, the Ga content of the electrode material must be less than 47%. By trying different values it is found that $Ga_{0.32}In_{0.67}As/Ga_{0.47}In_{0.53}As$ gives a conduction band offset of around 91meV. The lattice constant for $Ga_xIn_{1-x}As$ is given by $(6.0583 - 0.405x)$ Å. The lattice constants of these two materials are 5.9287 Å and 5.8680 Å, which correspond to just over 1% difference.

In order to decrease the work function of barriers to a value similar to the conduction band offset, the electrode material should be doped with a donor concentration of $2\times10^{23}$ m$^{-3}$. By introducing these extra carriers the electron quasi-Fermi level of both the electrode and barrier material is increased. Under zero bias the work function (measured between the chemical potential of the electrode and conduction band of the barrier material) is then around 90 meV. Because of the introduction of extra carriers band bending occurs and the maximum energy in the conduction band (or the minimum energy that carriers must overcome to contribute to thermionic current) increases. So even though the chemical potential is now higher than the conduction band of the electrode material, this increase in conduction band of the barrier material causes the work function to remain near the conduction band offset.

The barriers could also be doped with carriers (as in the design for third-generation devices) but by keeping the barrier material undoped the resistance of the barriers is much larger than the electrodes. This means that the majority of work (assuming constant current) is dissipated in the barriers themselves and not wasted in the electrodes, which would increase heat generation due to joule heating.

Because of the interplay between carrier concentration and electric field, the average work function height changes with applied bias. At zero bias it is around 90 meV whilst at a bias of 1 V the average barrier height decreases to around 67 meV.

### 7.5.4 Device Structure

As calculated in section 4.6.2 the mean free path of $Ga_{0.47}In_{0.53}As$ is around 93 nm so barriers should be around this width. If random barrier widths around this width are used then it is hoped that the average thermal conductivity of the barriers will be reduced due to phonon localisation. A range of barriers widths between 85-100 nm will hopefully promote this.

The width of electrodes is not as constrained as in the barriers. Ballistic motion is only

Table 7.1: $Ga_xIn_{1-x}As$ Material Parameters

| **Name** | **Value** | | **Unit** |
|---|---|---|---|
| $m_n$ | $0.023 + 0.037x + 0.003x^2$ | [79] | $m_0$ |
| $m_p$ | $0.41 + 0.1x$ | [79] | $m_0$ |
| $\mu_n$ | $(40 - 80.7x + 49.2x^2) \cdot 10^3$ | [79] | $cm^2/V{\cdot}s$ |
| $\mu_p$ | 350 | [79] | $cm^2/V{\cdot}s$ |
| $E_g$ | $0.42 - 4.19 \cdot 10^{-4} \cdot \frac{T^2}{T+271}$ $+ 0.625x - \left(\frac{5.8}{T+300} - \frac{4.19}{T+271}\right) \cdot 10^{-4} \cdot T^2 x$ $+ 0.475x^2$ | [79] | eV |
| $\chi$ | $4.731 - 0.42x - 0.24x^2$ | see Sec. 6.2.3 | eV |
| $\epsilon_r$ | $15.1 - 2.87x + 0.67x^2$ | [79] | $\epsilon_0$ |

required in the barriers so there is more freedom in varying the electrode widths in order to introduce disorder into phonon transport. A range of widths between 100-150 nm seems appropriate.

As with previous designs the number of barriers is set to 10. Increasing this number would increase the maximum cooling possible but also increases the difficulty in fabricating the devices. 20-50 barriers would theoretically increase the maximum cooling.

### 7.5.5 Substrate Selection

To ensure that minimal work is dissipated in the substrate it should be highly doped. n-type substrates with a carrier concentration of $5{\times}10^{18}$ $cm^{-3}$ are ideal for this application. One possible solution is to use highly doped GaAs for the substrate, although this introduces a lattice constant mismatch of around 4.5% with $Ga_{0.32}In_{0.67}As$.

### 7.5.6 Energy Band Diagram of Fourth-Generation Device

Fig. 7.11 shows the structure of a fourth-generation device. From section 4.6.2 the maximum cooling at room temperature is expected to be as much as 3.5 K, as opposed to only 0.57 K for third-generation devices.

```
Mat=GaInAs x=0.32 Nd=2e23  Width=500    Step=25
Mat=GaInAs x=0.32 Nd=2e23  Width=100    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=99     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=123    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=93     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=143    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=91     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=100    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=93     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=123    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=97     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=113    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=93     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=100    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=87     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=137    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=93     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=103    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=97     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=123    Step=5
Mat=GaInAs x=0.47 Nd=0     Width=93     Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=100    Step=5
Mat=GaInAs x=0.32 Nd=2e23  Width=500    Step=25
AppliedBias=0
```

Figure 7.11: Listing of ‘Ideal’ Fourth-Generation Device (no substrate included)

Fig. 7.12 shows the energy-band diagram of the proposed fourth-generation device under an applied bias of 0.5 V. At this bias all barrier heights are similar and equal to around 73 meV. This value is close to the value used to calculate maximum cooling (77 meV) in section 4.6.2. This design is superior to that of first- and second- generation devices because the electron quasi-Fermi level is above the conduction band of the electrodes due to doping in the electrodes.

By inspecting Fig. 7.12 it is seen that nearly all of the applied bias is dropped over the barrier regions. As discussed this is due to the barrier regions remaining undoped. This design, therefore, is advantageous over third-generation devices (see Fig. 7.7) which dissipate relatively more energy in the electrodes. This also reduces the effect that a highly doped substrate will have on device performance.

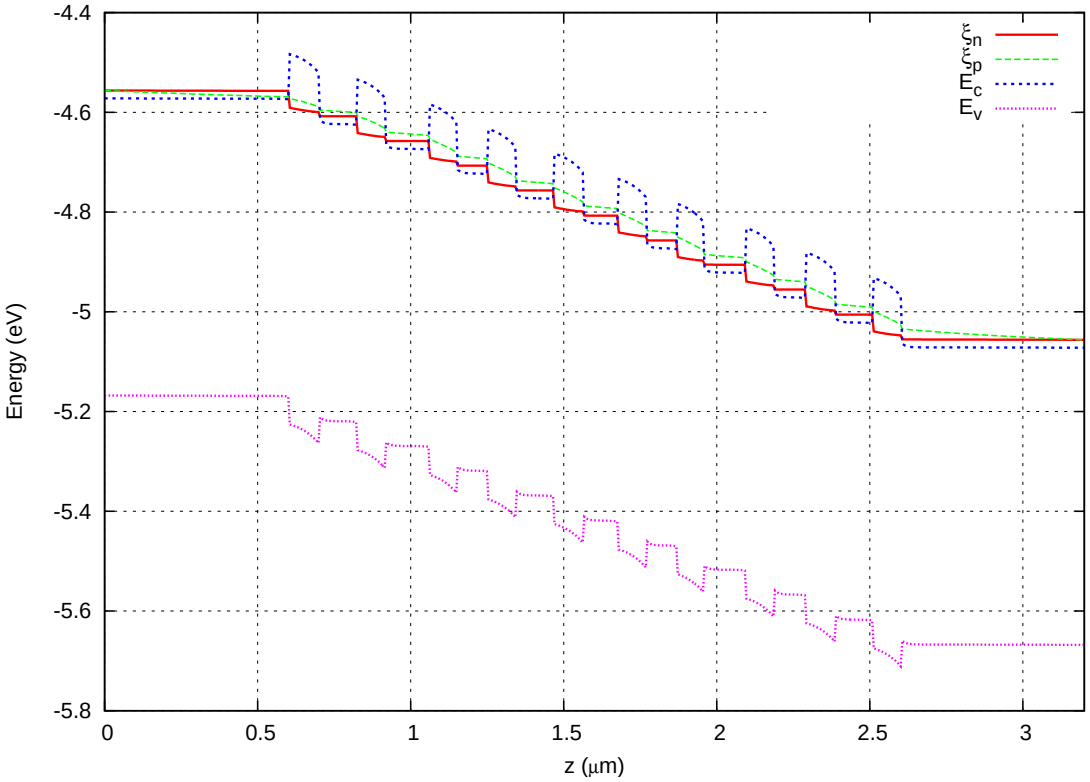


Figure 7.12: Energy-band diagram of ‘ideal’ device under applied bias 0.5 V.

### 7.5.7 Conclusions

By using the conclusions reached from earlier device designs a new device design has been proposed that has a number of distinct advantages. The most important is the use of a better barrier material, namely $Ga_{0.47}In_{0.53}As$. This immediately increases device performance by reducing heat conduction. Other design considerations, such as introducing random barrier and electrode widths in an attempt to localise phonons, aim to reduce heat conduction further.

By doping the electrodes and keeping the barriers undoped the barrier height of each barrier is reduced whilst still ensuring the majority of energy is dissipated in the barriers and heat dissipation is minimised in other parts of the device. This also has the advantage of reducing the effect that a highly doped substrate will have on device performance. Where cooling in previous generations may have been masked by internal heat generation

due to this dissipation, it is believed that substantial cooling should be observed in the next generation of devices.

# CHAPTER 8

# CONCLUSIONS

This thesis has investigated using GaAs/AlGaAs heterostructures and, more generally, semiconductor heterostructures to achieve thermionic cooling. Although inconclusive evidence was obtained to determine whether any cooling was actually achieved in the device fabricated in this work, the experimental electrical results did show that the model initially developed was inadequate, most likely due to the assumption that space charge accumulation did not significantly affect the devices. The simple models used in the first part of the thesis, although not sufficient in explaining experimental $I$-$V$ characteristics, do offer important information about single and multiple barrier devices. The theoretical results from these models should not be ignored. They do correctly predict some important trends in device behaviour. For instance, using different device parameters such as effective mass, mobility and thermal resistivity the models show which combination of parameters will theoretically produce the best devices. The heat transport equations show that the heat current will increase throughout the device as the work needed to transport the heat from one electrode to the next effectively becomes a heat source of its own which also must be removed in subsequent barrier jumps. This behaviour is not expected to change, even if a more detailed model is developed in which mechanisms such as internal heat generation, diffusive electrical transport and space-charge effects are included.

This single-barrier model in chapter 3 demonstrated that there are two classes of device. It was shown that Class 1 devices always perform better than Class 2 devices and should be used as they also simplify analysis. It is not expected that these results will differ with development of a more complete model.

The model developed in chapter 6 tries to address some of the incorrect assumptions made in the first model. The experimental $I$-$V$ characteristics are predicted much better by this new model. Energy transport has not yet been included in this model and it is suggested that this be considered for extension in the future. This model can, though, be used to design new generations of devices.

The lack of observable cooling could be attributed to a number of important factors that must be addressed. Firstly, the choice of material (GaAs) is definitely not the ideal. It was used in this work because of availability and because, at the time of commencing this work, it had never been tried for this type of device. The expected cooling, according to theory, was less than 1 K, even for an 'ideal' device. When non-ideal effects such as contact resistance, joule heating in the substrate, heating in the wire bonds, thermal mass of the micro-thermocouples, lattice damage and general packaging issues are taken in to account, it is not surprising that cooling was not observed. These effects should all be addressed, but the most noticeable improvement will be observed by using different material. Non-contact methods of temperature determination, such as optical methods, will also improve measurements as the thermal load of the surface will be significantly reduced.

It may be that due to the finite thermal resistivity of the material used, these types of thermionic devices may never perform sufficiently well that they can be used for low-temperature applications such as domestic refrigeration. They may still have uses in situ for temperature stabilisation of laser sources and microprocessors. Other, more promising types, such as opto-thermionic devices should be considered as they remove heat from the devices in the form of light rather than simply trying to transport the heat from one area to another whilst trying to inhibit conduction heat flow. Further advances in phonon localisation using superlattices may yet offer a solution to heat conduction in these devices.

# APPENDIX A

## PARTIAL DIFFERENTIAL EQUATIONS USED TO THE CALCULATE JACOBIAN MATRIX IN CHAPTER 4

## A.1 Calculation of the Jacobian Matrix

The Jacobian Matrix given by Eq. 4.12 is needed in order to update the system variables at each iteration of the Newton-Raphson method.The following sections show the partial differential equations of each of the system equations with respect to each of the system variables.

For each of the continuity equations only 5 entries in each row are non-zero. Careful variable ordering could be used in order to achieve a band-diagonal matrix. This would allow the use of special algorithms in order to invert the Jacobian but for this relatively simple system it was found unnecessary.

### A.1.1 Electrical Current Continuity Equations, $1 \leq i \leq (N-1)$

The first $N-1$ equations in the system are the electrical current continuity equations:

$$\begin{aligned} F_i &= J_i - J_{i+1} \\ &= AT_{i-1}^2 \exp\left(\frac{-q\phi_{i-1}}{k_B T_{i-1}}\right) - AT_i^2 \exp\left(\frac{-q\left(\phi_{i-1}+V_i\right)}{k_B T_i}\right) \\ &- AT_i^2 \exp\left(\frac{-q\phi_i}{k_B T_i}\right) + AT_{i+1}^2 \exp\left(\frac{-q\left(\phi_i+V_{i+1}\right)}{k_B T_{i+1}}\right) \qquad \text{(A.1)} \\ i &= 1, 2, \ldots, N-1 \end{aligned}$$

The partial differential equations with respect to system biases are given by:

$$\frac{\partial F_i}{\partial V_i} = \frac{qAT_i}{k_B} \exp\left[\frac{-q(\phi_{i-1}+V_i)}{k_B T_i}\right] \qquad \text{(A.2)}$$

$$\frac{\partial F_i}{\partial V_{i+1}} = -\frac{qAT_{i+1}}{k_B} \exp\left[\frac{-q(\phi_i+V_{i+1})}{k_B T_{i+1}}\right] \qquad \text{(A.3)}$$

$$\frac{\partial F_i}{\partial V_j} = 0 \text{ where } j \neq i, i+1 \qquad \text{(A.4)}$$

The partial differential equations with respect to system temperatures are given by:

$$\frac{\partial F_i}{\partial T_{i-1}} = \left(2AT_{i-1} + \frac{qA\phi_{i-1}}{k_B}\right)\exp\left[\frac{-q\phi_{i-1}}{k_B T_{i-1}}\right] \tag{A.5}$$

$$\begin{aligned}\frac{\partial F_i}{\partial T_i} = &-\left(2AT_i + \frac{qA(\phi_{i-1}+V_i)}{k_B}\right)\exp\left[\frac{-q(\phi_{i-1}+V_i)}{k_B T_i}\right] \\ &-\left(2AT_i + \frac{qA\phi_i}{k_B}\right)\exp\left[\frac{-q\phi_i}{k_B T_i}\right]\end{aligned} \tag{A.6}$$

$$\frac{\partial F_i}{\partial T_{i+1}} = \left(2AT_{i+1} + \frac{qA(\phi_i+V_{i+1})}{k_B}\right)\exp\left[\frac{-q(\phi_i+V_{i+1})}{k_B T_{i+1}}\right] \tag{A.7}$$

$$\frac{\partial F_i}{\partial T_j} = 0 \text{ where } j \neq i-1, i, i+1 \tag{A.8}$$

### A.1.2 Heat Current Continuity Equations, $N \leq i \leq 2(N-1)$

The second $N-1$ equations in the system are the heat current continuity equations:

$$\begin{aligned} F_i &= J_{qi}^{in} - J_{qi}^{out} \\ &= \left(\phi_{i-1} + V_i + \frac{2k_BT_{i-1}}{q}\right) AT_{i-1}^2 \exp\left(\frac{-q\phi_{i-1}}{k_BT_{i-1}}\right) \\ &- \left(\phi_{i-1} + V_i + \frac{2k_BT_i}{q}\right) AT_i^2 \exp\left(\frac{-q\left(\phi_{i-1}+V_i\right)}{k_BT_i}\right) \\ &- \frac{T_i - T_{i-1}}{R_i^{th}(N)} \\ &- \left(\phi_i + \frac{2k_BT_i}{q}\right) AT_i^2 \exp\left(\frac{-q\phi_i}{k_BT_i}\right) \\ &+ \left(\phi_i + \frac{2k_BT_{i+1}}{q}\right) AT_{i+1}^2 \exp\left(\frac{-q(\phi_i+V_{i+1})}{k_BT_{i+1}}\right) \\ &+ \frac{T_{i+1} - T_i}{R_{i+1}^{th}(N)} \qquad i = 1, 2, \ldots, N-1 \end{aligned} \tag{A.9}$$

The partial differential equations with respect to system biases are given by:

$$\begin{aligned} \frac{\partial F_i}{\partial V_i} &= AT_{i-1}^2 \exp\left[\frac{-q\phi_{i-1}}{k_BT_{i-1}}\right] \\ &+AT_i\left(\frac{q(\phi_{i-1}+V_i+2k_BT_i/q)}{k_B} - T_i\right)\exp\left[\frac{-q(\phi_{i-1}+V_i)}{k_BT_i}\right] \end{aligned} \tag{A.10}$$

$$\frac{\partial F_i}{\partial V_{i+1}} = -\left(\phi_i + \frac{2k_BT_{i+1}}{q}\right)\frac{qAT_{i+1}}{k_B}\exp\left[\frac{-q(\phi_i+V_{i+1})}{k_BT_{i+1}}\right] \tag{A.11}$$

$$\frac{\partial F_i}{\partial V_j} = 0 \text{ when } j \neq i, i+1 \tag{A.12}$$

The partial differential equations with respect to system temperatures are given by:

$$\frac{\partial F_i}{\partial T_{i-1}} = \left[\frac{2k_BT_{i-1}^2}{q} + \left(2T_{i-1} + \frac{q\phi_{i-1}}{k_B}\right)\left(\phi_{i-1} + V_i + \frac{2k_BT_{i-1}}{q}\right)\right] \times A\exp\left(\frac{-q\phi_{i-1}}{k_BT_{j-1}}\right) + \frac{1}{R_i^{th}(N)} \tag{A.13}$$

$$\begin{aligned}\frac{\partial F_i}{\partial T_i} = & -\left[\frac{2k_BT_i^2}{q} + \left(2T_i + \frac{q(\phi_{i-1}+V_i)}{k_B}\right)\left(\phi_{i-1} + V_i + \frac{2k_BT_i}{q}\right)\right] \\ & \times A\exp\left[\frac{-q(\phi_{i-1}+V_i)}{k_BT_i}\right] - \frac{1}{R_i^{th}(N)} \\ & -\left[\frac{2k_BT_i^2}{q} + \left(2T_i + \frac{q\phi_i}{k_B}\right)\left(\phi_i + \frac{2k_BT_i}{q}\right)\right] \\ & \times A\exp\left(\frac{-q\phi_i}{k_BT_i}\right) - \frac{1}{R_{i+1}^{th}(N)}\end{aligned} \tag{A.14}$$

$$\frac{\partial F_i}{\partial T_{i+1}} = \left[\frac{2k_BT_{i+1}^2}{q} + \left(2T_{i+1} + \frac{q(\phi_i+V_{i+1})}{k_B}\right)\left(\phi_i + \frac{2k_BT_{i+1}}{q}\right)\right] \times A\exp\left[\frac{-q(\phi_i+V_{i+1})}{k_BT_{i+1}}\right] + \frac{1}{R_{i+1}^{th}(N)} \tag{A.15}$$

$$\frac{\partial F_i}{\partial T_j} = 0 \text{ when } j \neq i-1, i, i+1 \tag{A.16}$$

# APPENDIX B

# THERMIONIC DEVICE SIMULATION INFORMATION

## B.1 User's Guide

*TISim* is a command-line program that can be compiled on many systems that have a decent C++ compiler.

### B.1.1 Compiling the Program

*TISim* can be compiled with G++ and GCC using the Makefile that comes with the source code. It can also be compiled with other compilers. Where available all optimisations should be enabled to increase execution time of the program. One very important option that must be used is to make sure no optimisations are performed on the source code that reduce floating-point precision. Under G++ this option is '--float-store'. For Microsoft C++ this option is '/Op'. If this option is not enabled then convergence may not be reached and if it is reached the results may not be correct.

### B.1.2 Running the Program

When started the following options will be made available to the user.

```
1. Load Device Structure File & Generate Device
2. Solve Equations & Output Data
3. Output Data
4. Change Bias of Right Electrode
5. Generate I-V Characteristics

6. Using Fermi-Dirac Statistics
7. Recombination Rates (SAO)
8. Including Thermionic Current

9. Show this menu
0. Exit

>:
```

Figure B.1: Menu Options of *TISim*

An explanation of the menu options in Fig. B.1 are as follows:

1. Load a device structure file from the ‘./DevFiles/’ directory immediately below the executable directory. See section 6.3.0.2 for an explanation of required input file format.

2. If the applied bias is set to zero (or not specified) in the device file only Poisson’s equation will be solved. If the applied bias is non-zero the three coupled electrical continuity equations (Poisson’s equation and the two current continuity equations) will be solved simultaneously. The output of this will be written to the file ‘Out.dat’ in the executable directory. The equations will be solved iteratively until either the error is less than the variable ‘CutOff’ (default set to $10^{-12}$) or the number of iterations exceeds ‘MaxIts’ (default value set to 20). Both of these variables can be changed in the constructor of the ‘Device’ class in the C++ file ‘Device.cpp’.

3. At any time the system state can be written to ‘Out.dat’.

4. Change the bias of right electrode.

5. The user will be prompted for a starting potential, end potential and potential increment of the right electrode. The system equations will be solved at each of these potentials and this potential vs. total current density (which should be near constant through the device for a given bias) will be written to the file ‘IV.dat’ in the executable directory.

6. Toggle between Fermi-Dirac and Maxwell-Boltzmann statistics. The default is to use Fermi-Dirac statistics. For non-degenerate materials, Maxwell-Boltzmann statistics can be used. The number of calculations needed will be reduced and convergence time will be increased.

7. Set which recombination-generation mechanisms will be included. In Fig. B.1 the mechanisms included are (S)chockley-Read-Hall, (A)uger and (O)ptical. The user can set which mechanisms are included by inputting a string with

'S', 'A', 'O' or 'I' in any order (case-sensitive). 'I' will include Impact Ionisation.

8. Toggle whether Thermionic Emission is included in the current density calculations.

9. Output these menu options at any time.

10. Exit *TISim*.

### B.1.3 Output file formats

Each column in the output files will be tab delimited.

#### B.1.3.1 'Out.dat'

'Out.dat' outputs the system state at each node in the device. A whole range of variables vs. distance at each node is output.

#### B.1.3.2 'IV.dat'

'IV.dat' is output when menu option '5.' is selected. Two columns will be output - $V$ (in volts) and $I$ which is the current density in $\mathrm{Am}^{-2}$.

## B.2 Block Diagram of Program

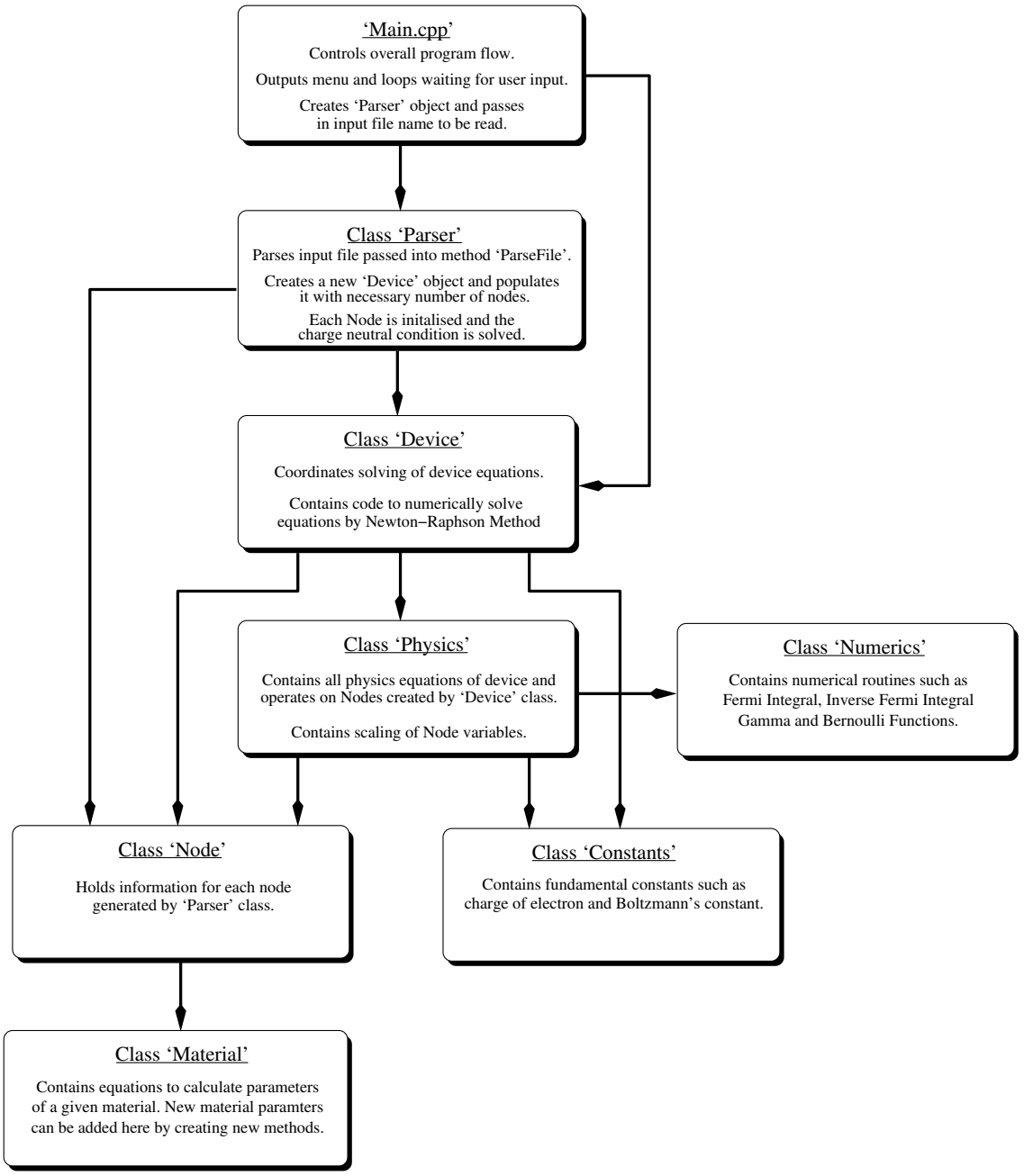


Figure B.2: Block diagram of *TISim*

*TISim* is written in object-oriented C++ and is a collection of objects. The purpose and relationship between each object is shown Fig. B.2.

## REFERENCES

[1] C. J. LaBounty, A. Shakouri, P. Abraham, and J. E. Bowers, "Integrated cooling for optoelectronic devices," in *Proc. SPIE*, vol. 3950, pp. 69–75, Apr 2000.

[2] L. Rushing, A. Shakouri, P. Abraham, and J. E. Bowers, "Micro thermoelectric coolers for integrated applications," in *16th International Conference on Thermoelectrics*, pp. 646–649, IEEE, 1997.

[3] G. D. Mahan, J. O. Sofo, and M. Bartkowiak, "Multilayer thermionic refrigerator and generator," *J. Appl. Phys.*, vol. 83, pp. 4683 – 4689, May 1998.

[4] H. J. Goldsmit, *Electronic Refrigeration*. London: Pion, 1986.

[5] G. D. Mahan, "Inhomogeneous thermoelectrics," *J. Appl. Phys.*, vol. 70, pp. 4551–4554, October 1991.

[6] B. Sherman, R. R. Heikes, and R. W. Ure, "Calculation of efficiency of thermoelectric devices," *J. Appl. Phys.*, vol. 31, pp. 1–16, January 1960.

[7] J. C. Bass and N. B. Elsner, "Segmented selenide thermoelectric generator.," in *Proceedings of the Third International Conference on Thermoelectric Energy Conversion*, (New York, NY, USA), pp. 8–12, IEEE, 1980.

[8] N. B. Elsner, J. Chin, and G. H. Reynolds, "Fabrication of selenide segmented elements.," in *Proceedings of the Third International Conference on Thermoelectric Energy Conversion*, (New York, NY, USA), pp. 105–108, IEEE, 1980.

[9] L. D. Hicks, T. C. Harman, X. Sun, and M. S. Dresselhaus, "Experimental study of the effect of quantum-well structures on the thermoelectric figure of merit," *Phys. Rev. B*, vol. 53, pp. R10493–R10496, April 1996.

[10] G. D. Mahan, "Thermionic refrigeration," *J. Appl. Phys.*, vol. 76, pp. 4362 – 4366, October 1994.

[11] A. Shakouri and J. E. Bowers, "Heterostructure integrated thermionic refrigeration," in *16th International Conference on Thermoelectrics*, pp. 636–640, IEEE, 1997.

[12] A. Shakouri and J. E. Bowers, "Heterostructure integrated thermionic coolers," *Appl. Phys. Lett.*, vol. 71, pp. 1234 – 1236, September 1997.

[13] G. D. Mahan and L. M. Woods, "Multilayer thermionic refrigeration," *Phys. Rev. Lett.*, vol. 80, pp. 4016 – 4019, May 1998.

[14] A. Shakouri, C. LaBounty, J. Piprek, P. Abraham, and J. E. Bowers, "Thermionic emission cooling in single barrier heterostructures," *Appl. Phys. Lett.*, vol. 74, pp. 88–89, January 1999.

[15] M. D. Ulrich, P. A. Barnes, and C. B. Vining, "Effect of contact resistance in solid-state thermionic refrigeration," *J. Appl. Phys.*, vol. 92, pp. 245–247, July 2002.

[16] E. T. Swartz, "Thermal resistance at interfaces," *Appl. Phys. Lett.*, vol. 51, pp. 2200–2202, December 1987.

[17] S.-M. Lee, D. G. Cahill, and R. Venkatasubramanian, "Thermal conductivity of Si–Ge superlattices," *Appl. Phys. Lett.*, vol. 70, pp. 2957–2959, June 1997.

[18] G. Chen and M. Neagu, "Thermal conductivity and heat transfer in superlattices," *Appl. Phys. Lett.*, vol. 71, pp. 2761–2763, November 1997.

[19] R. S. Prasher and P. E. Phelan, "A scattering-mediated acoustic mismatch model for the prediction of thermal boundary resistance," *J. Heat Transfer*, vol. 123, pp. 105–112, February 2001.

[20] T. Zeng and G. Chen, "Phonon heat conduction in thin films: Impacts of thermal boundary resistance and internal heat generation," *J. Heat Transfer*, vol. 123, pp. 340–347, April 2001.

[21] D. Huang, G. Gumbs, Y. Zhao, and G. W. Auner, "Optical-phonon transport and localization in periodic and Fibonacci polar-semiconductor superlattices," *Phys. Lett. A*, vol. 200, pp. 459–463, May 1995.

[22] G. Chen, T. Zeng, T. Borca-Tasciuc, and D. Song, "Phonon engineering in nanostructures for solid-state energy conversion," *Mat. Sci. Eng.*, vol. A292, pp. 155–161, 2000.

[23] R. Zhou, D. Dagel, and Y. H. Lo, "Multilayer thermionic cooler with a varying current density," *Appl. Phys. Lett.*, vol. 74, pp. 1767–1769, March 1999.

[24] A. G. Mul'shukov and K. A. Chao, "Opto-thermionic refrigeration in semiconductor heterostructures," *Phys. Rev. Lett.*, vol. 86, pp. 5570–5573, June 2001.

[25] Y. Hishinuma, T. H. Geballe, B. Y. Moyzhes, and T. W. Kenny, "Refrigeration by combined tunnelling and thermionic emission in vacuum: Use of nanometer scale design," *Appl. Phys. Lett.*, vol. 78, pp. 2572–2574, April 2001.

[26] Y. Hishinuma, B. Y. Moyzhes, T. H. Geballe, and T. W. Kenny, "Vacuum thermionic

refrigeration with a semiconductor heterojunction structure," *Appl. Phys. Lett.*, vol. 81, pp. 4242–4244, November 2002.

[27] X. C. Xuan, "Combined thermionic-thermoelectric refrigerator," *J. Appl. Phys.*, vol. 92, pp. 4746–4750, October 2002.

[28] G. S. Nolas and H. J. Goldsmid, "A comparison of projected thermoelectric and thermionic refrigerators," *J. Appl. Phys.*, vol. 85, pp. 4066–4070, April 1999.

[29] M. D. Ulrich, P. A. Barnes, and C. B. Vining, "Comparison of solid-state thermionic refrigeration with thermoelectric refrigeration," *J. Appl. Phys.*, vol. 90, pp. 1625–1631, August 2001.

[30] E. L. Chaffee, *Theory of thermionic vacuum tubes: fundamentals, amplifiers, detectors*. New York: McGraw-Hill, 1933.

[31] O. W. Richardson, *Emission of Electricity from Hot Bodies*. London, New York [etc.]: Longmans, Green and co., 5th ed., 1916.

[32] M. A. Omar, *Elementary Solid State Physics: Principles and Applications*. Reading, Mass.: Addison-Wesley Pub. Co., 1975.

[33] P. A. Tipler, *Physics for Scientists and Engineers*. Worth Pulishers, Inc., third ed., 1991. Extended Version.

[34] A. A. Grinberg, "Thermionic emission in heterosystems with different effective electron masses," *Phys. Rev. B*, vol. 33, pp. 7256–7258, May 1986.

[35] F. W. Sears and G. L. Salinger, *Thermodynamics, kinetic theory, and statistical thermodynamics*. Reading, Mass.: Addison-Wesley, 3rd ed., 1975.

[36] K. Seeger, *Semiconductor Physics: An Introduction*. Berlin; New York: Springer-Verlag, 4th ed., 1989.

[37] B. Sapoval and C. Hermann, *Physics of Semiconductors*. New York: Springer-Verlag New York, Inc., 1995.

[38] M. Shur, *Physics of Semiconductor Devices*. New Jersey: Prentice-Hall, Inc., 1990.

[39] L. Pfeiffer, K. W. West, H. L. Stormer, and K. W. Baldwin, "Electron mobilities exceeding $10^7$ $cm^2/V{\cdot}s$ in modulation-doped GaAs," *Appl. Phys. Lett.*, vol. 55, no. 18, pp. 1888 – 1890, 1989.

[40] M. S. Bresler, *Handbook Series on Semiconductor Parameters*, vol. 2: Ternary and Quaternary $A_3B_5$ Semiconductors, ch. 6, pp. 132 – 152. World Scientific Publishing Co. Pte. Ltd., 1999.

[41] G. Burns, *Solid State Physics*. Orlando, Fla.: Academic Press, 1985.

[42] Y. A. Goldberg, *Handbook Series on Semiconductor Parameters*, vol. 2: Ternary and Quaternary $A_3B_5$ Semiconductors. World Scientific Publishing Co. Pte. Ltd., 1999.

[43] G. Chen, "Thermal conductivity and ballistic-phonon transport in the cross-plane direction of superlattices," *Phys. Rev. B*, vol. 57, pp. 14958–14973, June 1998.

[44] B. Lough, S. P. Lee, R. A. Lewis, and C. Zhang, "Electronic thermal transport and thermionic cooling in semiconductor multi-quantum-well structures.," *Computer Physics Communications*, vol. 142, pp. 274 – 280, December 2001.

[45] R. L. Burden and J. D. Faires, *Numerical Analysis*. Boston: PWS-Kent Pub. Co., 5th ed., 1993.

[46] W. H. Press, *Numerical recipes in C: the art of scientific computing*. Cambridge; New York: Cambridge University Press, 2nd ed., 1992.

[47] F. J. Vesely, *Computational Physics – An Introduction*. New York: Plenum Press,

1994.

[48] S. Adachi, *GaAs and Related Materials: Bulk Semiconductors and Superlattice Properties*. Singapore: World Scientific Publishing Co. Pte. Ltd., 1994.

[49] S. P. Lee, "Multilayer thermionic cooling in GaAs-$Al_xGa_{1-x}As$ heterostructures," Master's thesis, University of Wollongong, 2002.

[50] M. Sydor, J. Angelo, J. J. Wilson, W. C. Mitchel, and M. Y. Yen, "Photoreflectance from GaAs and GaAs/GaAs interfaces," *Phys. Rev. B*, vol. 40, pp. 8473–8484, October 1989.

[51] P. Kraisingdecha and M. Gal, "Differential reflectance spectroscopy of GaAs/GaAlAs at elevated temperatures," *Appl. Phys. Lett.*, vol. 69, pp. 1355–1357, September 1996.

[52] N. Dai, J. F. Feng, L. Y. Chen, X. Q. Liu, W. Lu, and J. L. Zhong, "Differential reflectance spectroscopy of GaAlAs thin films and GaAs bulk under externally applied temperature gradient," *Appl. Phys. Lett.*, vol. 70, pp. 2271–2273, April 1997.

[53] B. C. C. Lough, S. P. Lee, Z. Dou, R. A. Lewis, and C. Zhang, "Investigation into space charge effects in I-V characteristics of multi-layer semiconductor thermionic devices," *Physica E*, vol. 17, pp. 651 – 653, 2003.

[54] J. B. Scott, "Extension of Langmuir space-charge theory into the accelerating field range," *J. Appl. Phys*, vol. 52, p. 4406, July 1981.

[55] D. J. Griffiths, *Introduction to Electrodynamics*. Prentice Hall, 1999.

[56] H. K. Gummel, "A self-consistent iterative scheme for one-dimensional steady state transistor calculations," *IEEE Trans.*, pp. 455–465, 1964. ED-11.

[57] J. W. Slotboom, “Iterative scheme for 1- and 2- dimensional d.c.-transistor simulation,” *Electronics Letters*, vol. 5, pp. 677–678, December 1969.

[58] A. H. Marshak and K. M. V. Vliet, “Current in semiconductors with position-dependent band gap and electron affinity,” *Phys. Stat. Sol.*, vol. 42, pp. 279–283, 1977.

[59] C. C. McAndrew and K. Singhal, “A consistent nonisothermal extension of the Scharfetter-Gummel stable difference approximation,” *IEEE Elec. Dev. Lett.*, vol. EDL-6, pp. 446–447, September 1985.

[60] C. T. Wang, “A new set of semiconductor equations for computer simulation of submicron devices,” *Solid-State Electronics*, vol. 28, no. 8, pp. 783–787, 1985.

[61] C. M. Snowden, *Semiconductor Device Modelling*. IEE materials and devices, London, U. K.: P. Peregrinus on behalf of the Institution of Electrical Engineers, 1988.

[62] P. Roblin and H. Rohdin, *High-speed heterostructure devices*. Cambridge: Cambridge University Press, 2002.

[63] A. Trellakis, A. J. Galick, and U. Ravaioli, “Rational Chebyshev approximation for the Fermi-Dirac integral,” *Solid-State Electronics*, vol. 41, pp. 771 – 773, May 1997.

[64] P. Van Halen and D. L. Pulfrey, “Accurate, short series approximations to Fermi-Dirac integrals of order -1/2, 1/2, 1, 3/2, 2, 5/2, 3 and 7/2,” *J. Appl. Phys.*, vol. 57, pp. 5271–5274, June 1985.

[65] P. Van Halen and D. L. Pulfrey, “Erratum: Accurate, short series approximations to Fermi-Dirac integrals of order -1/2, 1/2, 1, 3/2, 2, 5/2, 3 and 7/2,” *J. Appl. Phys.*, vol. 59, pp. 2264–2265, March 1986.

[66] X. Aymerich-Humet, F. Serra-Mestres, and J. Millan, “A generalized approximation of the Fermi-Dirac integrals,” *J. Appl. Phys.*, vol. 54, pp. 2850–2851, May 1983.

[67] M. Kurata, *Numerical Analysis for semiconductor devices.* D. C. Heath and Company, 1982.

[68] A. De Mari, *Accurate numerical steady-state and transient one-dimensional solutions of semiconductor devices.* PhD thesis, California Institute of Technology Division of Eng. and Appl. Science, October 1967.

[69] R. L. Anderson, "Experiments on Ga-GaAs heterojunctions," *Solid-State Electron.*, vol. 5, pp. 341–344, September–October 1962.

[70] R. Dingle, W. Wiegmann, and C. H. Henry, "Quantum states of confined carriers in very thin $Al_xGa_{1-x}As$-GaAs-$Al_xGa_{1-x}As$ heterostructures," *Phys. Rev. Lett.*, vol. 33, pp. 827–830, September 1974.

[71] M. J. Kelly, *Low-Dimensional Semiconductors: Material, Physics, Technology, Devices*. Oxford: Clarendon Press; New York: Oxford University Press, 1995.

[72] R. Stratton, "Semiconductor current-flow equations (diffusion and degeneracy)," *IEEE Transactions on Electron Devices*, vol. ED-19, pp. 1288 – 1292, December 1972.

[73] Y. Apanovich, P. Blakey, R. Cottle, E. Lyumkis, B. Polsky, A. Shur, and A. Tcherniaev, "Numerical simulation of submicrometer devices including couple nonlocal transport and nonisothermal effects," *IEEE. Tran. Elec. Dev.*, vol. 42, pp. 890–898, May 1995.

[74] A. W. Smith and A. Rohatgi, "Non-isothermal extension of the Scharfetter-Gummel technique for hot carrier transport in heterostructure simulations," *IEEE. Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 12, pp. 1515–1523, October 1993.

[75] K. Yang, J. East, and G. Hadda, "Numerical modelling of abrupt heterojunctions

using a thermionic boundary condition," *Solid State Electronics*, vol. 36, pp. 321–330, 1993.

[76] K. Horio and H. Yanai, "Numerical modelling of heterojunctions including the thermionic emission mechanism at the heterojunction interface," *IEEE Trans. Electron Devices*, vol. 37, pp. 1093–1098, 1990.

[77] C. M. Snowden, *Introduction to Semiconductor Device Modelling*. Singapore: World Scientific Publishing Co Pte Ltd, 1986.

[78] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon read diode oscillator," *IEEE Transactions on Electron Devices*, vol. ED-16, pp. 64–77, January 1969.

[79] Y. A. Goldberg and N. M. Shmidt, *Handbook Series on Semiconductor Parameters*, vol. 2: Ternary and Quaternary $A_3B_5$ Semiconductors. World Scientific Publishing Co. Pte. Ltd., 1999.