Analytical Modelling of Isotype Heterojunctions

ANALYTICAL MODELLING OF ISOTYPE HETEROJUNCTIONS

BY

MANUEL GIL, M.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ENGINEERING PHYSICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2013) McMaster University

(Engineering Physics) Hamilton, Ontario, Canada

TITLE:	Analytical Modelling of Isotype Heterojunctions
AUTHOR:	Manuel Gil M. Sc. (Mathematics), Queen's University, Kingston, ON, Canada, B. Sc. (Applied Mathematics), Western University, London, ON, Canada.
SUPERVISOR:	Dr. Rafael Kleiman
NUMBER OF PAGES:	ix, 55

Abstract

An *isotype heterojunction* is a junction between two layers of dissimilar semiconductors both of which are doped either n-type or p-type. These semiconductor structures are found in a variety of optoelectronic devices, such as solar cells, semiconductor lasers, and detectors. Motivated by the structure of third generation inorganic solar cells, this thesis concentrates on the analytical modelling of isotype heterojunctions and its application to the design optimization of these devices. The main development of this work is the introduction of an analytical expression for the current density across an isotype heterojunction valid for arbitrary doping concentration ratios. This result generalizes the standard expression found in the literature, which is limited by the assumption that the doping concentration ratio between the two sides of the heterojunction is equal to one. The generalization is developed by employing the Lambert W function in the solution of the electrostatic boundary condition associated with the heterojunction interface. As done in the derivation of the standard expression found in the literature, the generalization only considers thermionic emission, but the same method can readily be applied for other transport mechanisms. A key feature of this generalized result is that it mathematically contains the expression for the current density across a metal-semiconductor Schottky contact as a limiting case, thereby unifying the treatment of these two heterointerfaces into a single general analytical description. This

latter find is particularly significant from a theoretical perspective, considering that the two heterointerfaces are traditionally described as separate topics in the presentation of semiconductor device theory.

Acknowledgments

I would firstly like to express my gratitude towards my supervisor, Dr. Rafael Kleiman, for accepting me as his graduate student and for all his valuable guidance and advice during the completion of this work. I am also very grateful to Jingfeng Yang for all his insight and valuable input regarding this research.

I also want to thank the School of Graduate Studies at McMaster University, the Department of Engineering Physics, and the Natural Sciences and Engineering Research Council (NSERC) of Canada for their funding; the staff in the Department of Engineering Physics for all their assistance with practical matters; and all graduate students who have in one way or another been supportive and instructive, especially Abhi Rampal, Martin Gerber, and Kevin Boyd.

I am always immensely grateful to my parents for their unconditional support.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

While the emphasis of this thesis is on the analytical study of a particular topic in the physics of semiconductor heterojunctions, the problem has been motivated within the context of multijunction solar cell design.

There is no question that developing efficient and readily available renewable energy technologies is one of the main priorities of current science and engineering research. The prospect of higher efficiency and lower production costs for solar cells has placed the field of solar photovoltaics as one of the most promising research avenues to tackle this issue, as evidenced by the emergence of a variety of commercially feasible solar cell technologies leading to the global cumulative solar photovoltaic capacity surpassing the 100 GW installed electrical power benchmark in 2012 (European Photovoltaic Industry Association (EPIA), 2013).

The different solar cell technologies tackle a variety of design goals and application

needs (Fonash, 2010; Nelson, 2003). In the case of multijunction solar cells, the strategy is to maximize efficiency by stacking individual solar cells in tandem. The cells form a sequence of decreasing band gaps in the direction of light travel, connected via tunnel junctions, and with the band gaps spread so as to cover as much as possible of the solar spectrum. Selecting the appropriate band gap values, this design improves cell efficiencies by expanding the energy range where photons can be absorbed from the solar spectrum, while at the same time minimizing thermalization losses. These solar cells have reached a 44.4% efficiency mark under 302 suns (Sharp Corporation, 2013) and recent work on multijunction device numerical modelling and optimization has led to 3-junction designs expected to break the 50% benchmark (Leite *et al.*, 2013). A typical multijunction solar cell is illustrated in Figure 1.1.

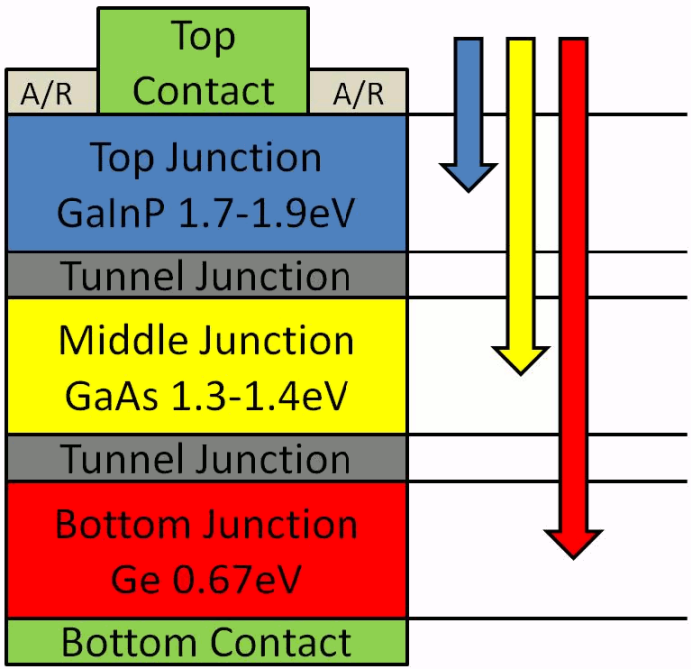

Figure 1.1: A typical example of a multijunction solar cell. Individual cells are connected in tandem through tunnel junctions, arrayed so that they form a sequence of decreasing band gaps in the direction of light travel through the cell (Weisse, 2010).

The standard set up for commercially available multi-junction cells is to use high band gap materials like InGaP and InGaAs grown on germanium substrates, which are much more expensive than silicon (Si) substrates. While very high efficiency may be the dominant design goal in some niche applications, in the context of large scale commercialization, a more important metric is the overall device cost to efficiency ratio. Therefore, an interesting alternative to consider is the development of high-efficiency silicon substrate multijunction cells. Figure 1.2 illustrates this concept for a double junction device. Note that SC stands for an unspecified relevant semiconductor, which would be determined according to an efficiency/cost optimization process and other design variables. One important design consideration that arises here is the choice of location for the tunnel junction, i.e., within the Si (n++Si/p++Si), within the SC (n++SC/p++SC), or as a hetero tunnel junction between Si and SC (n++Si/p++SC), leading to the following possibilities:

- Si tunnel junction on p-Si substrate: p-Si/p++Si/**n++Si/n-SC**,
- Si tunnel junction on n-Si substrate: n-Si/n++Si/**p++Si/p-SC**,
- SC tunnel junction on p-Si substrate: **p-Si/p++SC**/n++SC/n-SC,
- SC tunnel junction on n-Si substrate: **n-Si/n++SC**/p++SC/p-SC
- hetero tunnel junction on p-Si substrate: p-Si/p++Si/n++SC/n-SC,
- hetero tunnel junction on n-Si substrate: n-Si/n++Si/p++SC/p-SC.

Notice that in the structures that do not use the hetero tunnel junction, a heterojunction is formed between Si and SC where both semiconductors have the same

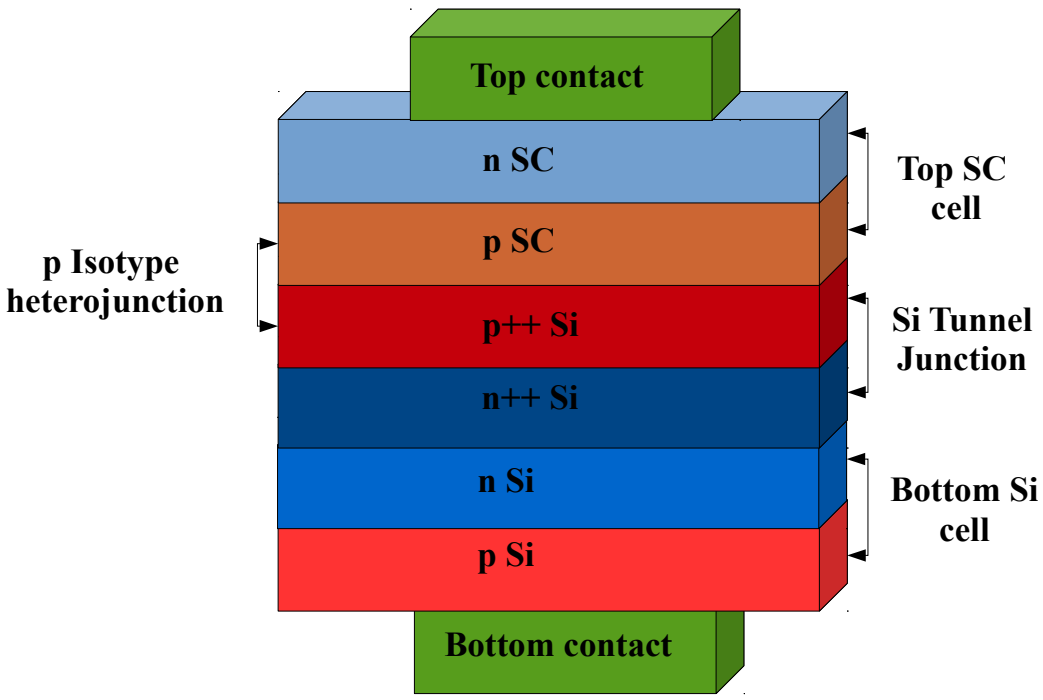

Figure 1.2: An example of a double junction Si-substrate solar cell where SC is a relevant semiconductor.

doping type, i.e., the highlighted elements **p++Si/p-SC** and **n-Si/n++SC**. This thesis focuses on the analytical study of the I-V characteristics of these *isotype heterojunctions*.

Section 1.2 introduces previous relevant work on the analytical modelling of isotype heterojunctions, which, to the best knowledge of the author, has until now relied on the assumption that the doping concentration ratio between the two sides of the heterojunction is equal to one. The generalization of the existing result developed in this thesis requires the use of the Lambert W function, which is not yet as well known as other standard special functions of mathematical physics. A brief introduction to this

function is presented in Chapter 2. This is followed by the full derivation of the generalized results in Chapter 3, where it is also demonstrated that the work developed here effectively unifies the treatment of isotype heterojunctions and metal-semiconductor Schottky contacts into a single analytical framework. Chapter 4 returns to the original motivation of this thesis and considers the behaviour of the differential resistance of a general isotype heterojunction for low applied voltages. Lastly, concluding remarks are presented in Chapter 5.

1.2 Existing Results in the Analytical Modelling of Isotype Heterojunctions

An *isotype heterojunction* is a junction between two layers of dissimilar semiconductors both of which are doped either n-type or p-type. The analysis of the current-voltage characteristic of these structures has led to the development of an energy band analytical model (Sze and Ng, 2007; Chang, 1965; Anderson, 1962), which was introduced by Anderson (Anderson, 1962). Anderson's model is formulated in analogy with the emission theory for metal-semiconductor diodes. This model is based on a band diagram with the general form given in Figure 1.3, where the specific barrier height and width varies according to the material combination being considered.

Because of the form of the band diagram in Figure 1.3, the dominant current transport mechanism is assumed to be thermionic emission. Anderson's model also neglects quantum mechanical transmission and reflection, diffusion across the barrier, bulk resistance, carrier recombination, interface states, minority carrier transport, as well as

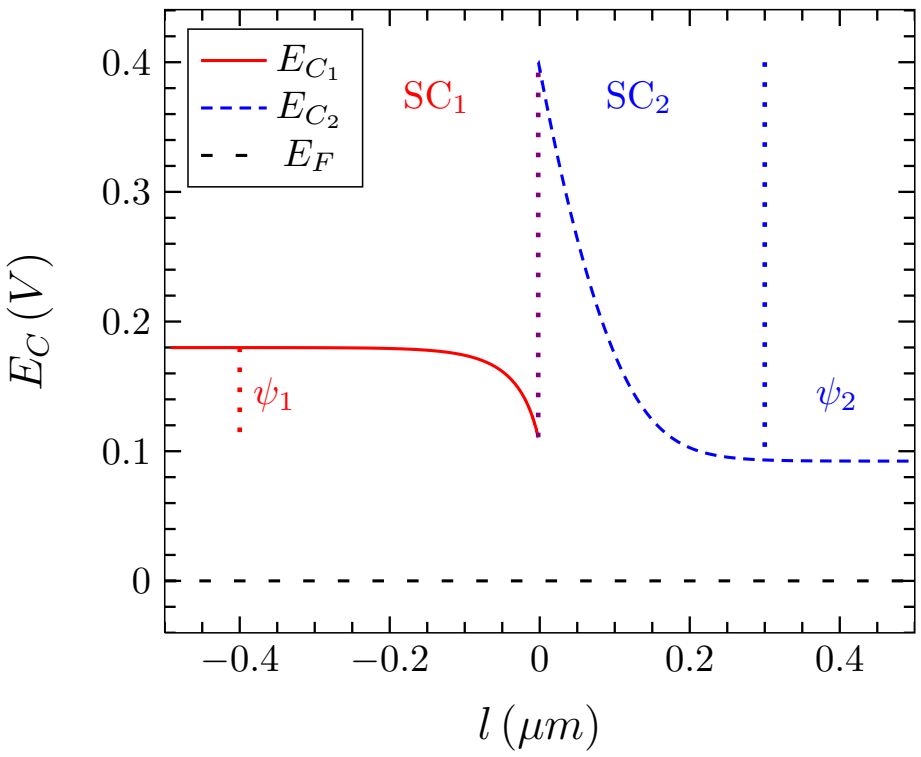

Figure 1.3: Typical n-n heterojunction conduction band profile between two dissimilar semiconductors SC$_1$ and SC$_2$, with associated conduction band energies, E_{C_1} and E_{C_2}, respectively, and a Fermi level E_F.

image force lowering (Sze and Ng, 2007). Anderson's original derivation and the works that followed (Sze and Ng, 2007; Chang, 1965; Anderson, 1962) also rely on the assumptions that $\alpha \equiv \epsilon_1 N_{D1}/\epsilon_2 N_{D2} = 1$, where N_{Di} and ϵ_i are the dopant concentration and the electric permittivity of the semiconductor region SC$_i$, respectively, and that $\psi \equiv \psi_1 + \psi_2 \gg kT/q$, where q is the electron charge, k is the Boltzmann constant, and T is the temperature. With these assumptions one can obtain the following expression for the current density $J(V)$:

$$J_1(V) = J_0 \frac{q\psi}{kT} e^{-q\psi/kT} \left(1 - \frac{V}{\psi}\right) \left[e^{qV/kT} - 1\right] , \tag{1.1}$$

where $J_0 \equiv qN_{D2}\sqrt{kT/2\pi m_2^*}$, m_2^* is the effective mass in semiconductor SC_2, and V is the applied voltage. Note that Equation 1.1 is qualitatively similar to the current density for a diode.

The assumption that $\alpha \equiv \epsilon_1 N_{D1}/\epsilon_2 N_{D2} = 1$ is a major limitation in the derivation of Equation 1.1 and its applicability in the context of device modeling, since doping concentration levels are one of the key controllable design parameters in semiconductor devices. Moreover, in Chapter 3 of this work, it will be demonstrated that the assumption $\alpha = 1$ is not supported by a physically meaningful observation, and is justified only from the point of view of significantly simplifying the underlying mathematics. More specifically, the $\alpha = 1$ assumption allows the avoidance of a transcendental equation that results from the semiconductor electrostatic boundary conditions. Fortunately however, such an equation belongs to a class of linear transcendental equations that are solvable via the Lambert W function, introduced in Chapter 2. A generalization of Equation 1.1 allowing for arbitrary values of α is derived in Chapter 3.

Chapter 2

The Lambert W Function

The Lambert W function is the major mathematical tool employed in the generalization presented in this work. While it has recently become another important special function in the toolkit of mathematical physics, its relatively recent rediscovery makes it still somewhat unknown in many circles of physics and applied mathematics. For this reason, and because of its crucial importance in the generalization developed by this thesis, this chapter provides a short introduction to the Lambert W function and a brief reference to some of its recent applications to physics. Unless otherwise specified, the material in this chapter is found in the seminal work by Corless *et. al* (Corless *et al.*, 1996), which is an excellent reference for a thorough introduction to this function and many of its applications.

2.1 A Brief History

The origin of the Lambert W function can be traced back to 1758 in relation to the work of Lambert on the solution of trinomial equations of the form $x = q + x^m$. Lambert

provided a solution to these equations in terms of a series in powers of q. Later work by Euler on a symmetric version of the equations and their special cases led to considering equations of the form $\log x = vx^{\alpha}$, the solution of which can be written in terms of the series

$$\log x = v + \frac{2^1}{2!}v^2 + \frac{3^2}{3!}v^3 + \frac{4^3}{4!}v^4 ...$$

The series converges for $|v| < 1/e$ and defines what is now known as the tree function $T(v)$, which satisfies $T(v) = -W(-v)$, where W is the Lambert W function defined in Section 2.2. The name inherits the term "Lambert" due to its connection to the work of Lambert on trinomial equations, and the letter W refers to the work of E. M. Wright (see Corless *et al.* (1996) and references there in). In this document, the terms 'W', 'W function' and 'Lambert W function', are all interchangeable.

2.2 Definition and Basic Properties

Given a map $f : \mathbb{C} \to \mathbb{C}$ defined by $z \mapsto ze^z$, the Lambert W function, $W(z)$, is defined as the *multivalued inverse* of f, i.e.,

$$W(z)\exp(W(z)) = z\,, \quad \forall z \in \mathbb{C}\,. \tag{2.1}$$

The defining equation always has an infinite number of solutions in $\mathbb{C}$, given by the complex branches of the Lambert W function, denoted by W_j, where $j \in \mathbb{Z}$. If $z \in \mathbb{R}$ there can be at most two real solutions, corresponding to the branches W_0 (principal branch) and W_{-1}, which are plotted in Figure 2.1.

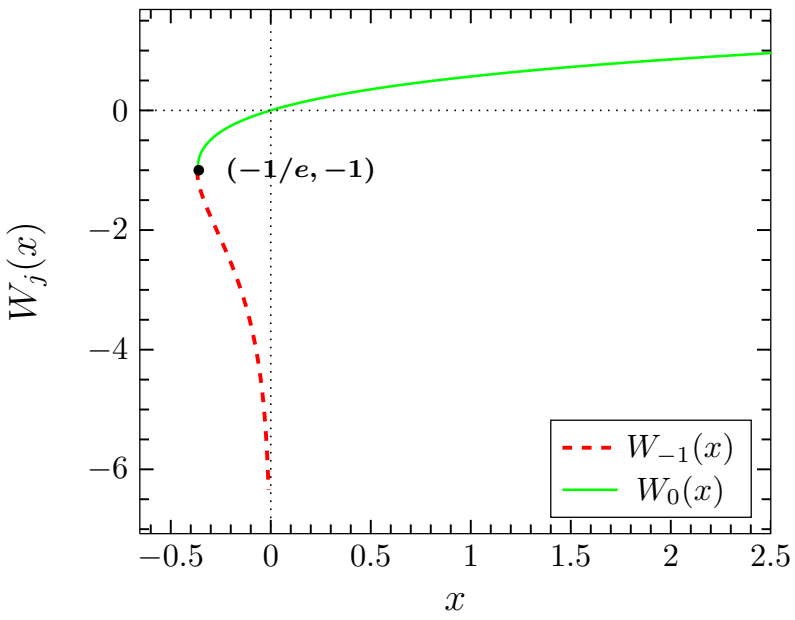

Figure 2.1: Real branches of the Lambert W function.

The solid line represents W_0, which is defined on $x \in [-1/e, \infty)$ with range $[-1, \infty)$, while the dashed line represents W_{-1}, defined on $x \in [-1/e, 0)$ with range $(-\infty, -1]$. An important observation is that *real values of $W(x)$ exist only for $x \geq -1/e$, and unique real values exist only for $x > 0$.* The possibility of a double solution (or even infinite when the domain of applicability extends to $\mathbb{C}$) can have important consequences in applications where solutions to equations are expressed in terms of W. In particular, we can see from Figure 2.1 that the behaviour of the two real branches is quite different, so that special care must be taken when arguing for the selection of one branch versus the other.

2.3 Some Useful Mathematical Properties

From the point of view of applications, the calculus properties of W are especially important. Applying implicit differentiation to Equation 2.1 one obtains an expression for the derivative of $W(x)$ that is implicitly defined in terms of $W(x)$:

$$\frac{\mathrm{d}W}{\mathrm{d}x} = \frac{W(x)}{x(1+W(x))} \qquad x \notin \{0, -1/e\}\,, \tag{2.2}$$

where the restriction at $x = -1/e$ corresponds to the location of the branch point joining W_0 and W_{-1} (See Figure 2.1). More generally, the n-th order derivatives of $W(x)$ are given as

$$\frac{d^n W(x)}{dx^n} = \frac{\exp(-nW(x))\, p_n(W(x))}{(1+W(x))^{2n-1}}\,, \tag{2.3}$$

where the polynomials $p_n(w)$ are defined via the recursive relation

$$p_{n+1}(w) = -(nw + 3n - 1)p_n(w) + (1+w)p_n'(w)\,,$$

with $p_1(w) = 1$ and $p_n(0) = (-n)^{n-1}/n!$. Note that it is also possible to obtain another expansion for $n-$th order derivatives in terms of second order Eulerian numbers. The antiderivative of $W(x)$ can also be obtained implicitly in terms of $W(x)$:

$$\int W(x)\,\mathrm{d}x = e^{W(x)}\left(W^2(x) - W(x) + 1\right) + C = x\left(W(x) - 1 + \frac{1}{W(x)}\right) + C. \tag{2.4}$$

Equation 2.4 follows from the definition of W and the observation that

$$\int f^{-1}(x)dx = yf(y) - \int f(y)dy\,,$$

where $x = f(y) \Leftrightarrow f^{-1}(x) = y$. Another important and useful observation is that the principal branch $W_0(x)$ is analytic at $x = 0$, with the series

$$W_0(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!}\, x^n = x - x^2 + \frac{3}{2}x^3 - \frac{8}{3}x^4 + \cdots\,, \tag{2.5}$$

which has a radius of convergence of $1/e$. Finally, one can obtain asymptotic expansions for $W_0(x)$ and W_{-1}, useful for large x and for negative x near zero, respectively, in terms of the expression

$$\begin{aligned} A(a,b) &= a - b + \sum_k \sum_m C_{km} a^{-k-m-1} b^{m+1} \\ &= a - b + \frac{b}{a} + \frac{b(-2+b)}{2a^2} + \frac{b(6-9b+2b^2)}{6a^3} + \ldots. \end{aligned} \tag{2.6}$$

such that $W_0 = A(\ln x, \ln(\ln x))$ and $W_{-1} = A(\ln(-x), \ln(-\ln(-x)))$, where the coefficients C_{km} are given in terms of the Stirling numbers of the second kind (Veberic, 2010; Corless *et al.*, 1996).

2.4 An Example of How to Use the Lambert W Function

Many transcendental equations can be solved via the Lambert W function, provided one can transform the equation to the form $xe^x = p$, where x is the variable of interest (or a particular function of the variable of interest) and p a quantity independent of x.

Once in this form, the definition of W yields the solution $x = W(p)$, where again, the number of solutions can vary from zero to infinitely many, depending on whether one defines the solution space over $\mathbb{C}$, $\mathbb{R}$, or any relevant subset.

As an example, suppose we wish to solve for x in the equation

$$e^{ax} + bx + c = 0 \quad a, b \neq 0\,. \tag{2.7}$$

With a few algebraic manipulations we can transform it to

$$\frac{a}{b} e^{-ac/b} = -\left(ax + \frac{ac}{b}\right) e^{-(ax+ac/b)},$$

which is now in the canonical form. Applying W to both sides and solving for x we obtain

$$x = -\frac{W\left(\frac{a}{b} e^{-ac/b}\right) + \frac{ac}{b}}{a}\,. \tag{2.8}$$

Other forms of transcendental equations and their solutions in terms of W are presented in Table 2.1.

2.5 Applications to Physics and Engineering

Since its rediscovery in the 1990s, the Lambert W function has been applied to a myriad of problems in many different areas in physics, computer science, as well as applied and pure mathematics. Corless *et al.* (1996) and references therein discuss applications to jet fuel problems, combustion, enzyme kinetics, as well as the resolution of a paradox in the calculation of the exchange forces between two nuclei within the hydrogen

molecule, among others. In addition, the Lambert W function has been recently applied to problems in classical and quantum statistical mechanics (Caillol, 2003; Valluri *et al.*, 2009; Tanguay *et al.*, 2010; Chandrashekar and Segar, 2013; Starikov, 2010), gravitation (Farrugia *et al.*, 2007; Mann and Ohta, 1997; Germani and Liguori, 2009), information theory (Jizba and Arimitsu, 2006), exact solutions to a type of Schrödinger equation (Williams, 2005), chemical physics (Kast and Tomazic, 2012), classical mechanics (Morales, 2011), electromagnetics (Valluri *et al.*, 2000; Jenn, 2002), thermoelectrics (Molli *et al.*, 2011), Fokker-Planck equations in the small noise limit (Lutz, 2005), and solar cells (Jain and Kapoor, 2005a,b; Jain *et al.*, 2006).

The Lambert W function has been implemented on a number of commercially available scientific computing packages such as Maple, Matlab, and Mathematica, as well as in the free, open-source, GNU Scientific Library (GSL), making the numerical evaluation of the different branches a routine matter.

Table 2.1: Some transcendental equations solved by the Lambert W function.

Equation	**Solution**
$x b^x = a$	$x = \dfrac{W(a \log b)}{\log b}$
$x^{dx^a} = b$	$x = \exp\left[\dfrac{1}{a} W\left(\dfrac{a}{d} \log b\right)\right]$
$h(z) = z^{z^{z^{z^{\cdot^{\cdot}}}}}$	$h(z) = -\dfrac{W(-\log z)}{\log z}$, for $z \in [e^{-e}, e^{\frac{1}{e}}]$
$\log(a + by) + cy = \log d$	$y = \dfrac{1}{c} W\left(\left[\dfrac{cd}{b}\right] \exp\left[\dfrac{ac}{b}\right]\right) - \dfrac{a}{b}$, for $a \geq 0; b, d > 0$
$p^{ax+b} = cx + d$	$x = -\dfrac{W(-\frac{a \ln p}{c} p^{b - \frac{ad}{c}})}{a \ln p} - \dfrac{d}{c}$, for $p > 0$ and $c, a \neq 0$

Chapter 3

An Analytical Expression for Thermionic Transport through Isotype Heterojunction Interfaces of Arbitrary Doping Ratio

This chapter generalizes the analytical expression for the current density across an isotype heterojunction, allowing for arbitrary doping ratios between the two sides of the heterojunction. By letting the parameter $\alpha \equiv \epsilon_1 N_{D1}/\epsilon_2 N_{D2}$ vary over all positive real numbers, the new expression significantly expands the space of isotype heterojunctions that can be analytically described, thereby improving the modelling scope for a variety of opto-electronic devices that rely on isotype heterojunctions, such as semiconductor lasers, detectors, and solar cells (Milnes and Feucht, 1972; Casey and Panish, 1978; Nelson, 2003). Besides the practical implications of the generalization, this new result

is also of theoretical significance. As shown in Section 3.2.1, the generalized result contains the expression for the current density across a metal-semiconductor Schottky contact as the special case $\alpha \to \infty$ under a fixed barrier height ψ_2 in the semiconductor region SC_2. It is important to note that isotype semiconductor heterojunctions and metal-semiconductor Schottky contacts have been traditionally described independently (e.g. Sze and Ng (2007)).

3.1 Derivation of the Generalized Current Density

The generalization starts by following the same physical arguments that Anderson provided in his original model. Recall that the relevant band diagram is assumed to be of the form given in Figure 1.3. Assuming Maxwell-Boltzmann statistics for $l < 0$ and a depletion region for $l > 0$, the electric displacement at the interface, $l = 0$, is

$$\mathscr{D} = \begin{cases} \sqrt{2\epsilon_1 N_{D_1} \left(kT \left[e^{U_1/kT} - 1 \right] - U_1 \right)}, & l \to 0^- , \\ \sqrt{2\epsilon_2 N_{D_2} U_2}, & l \to 0^+ . \end{cases} \tag{3.1}$$

where $U_i \equiv q(\psi_i - V_i)$, V_i and ψ_i are the voltage drop and the barrier height in SC_i, respectively, q is the electron charge, k is the Boltzmann constant, and T is the temperature (Anderson, 1962; Sze and Ng, 2007). Requiring the continuity of the electric displacement in Equation 3.1 at the interface imposes a relationship between V_1 and V_2, namely

$$\epsilon_1 N_{D_1} \left(kT \left[e^{q(\psi_1 - V_1)/kT} - 1 \right] - q(\psi_1 - V_1) \right) = \epsilon_2 N_{D_2} q(\psi_2 - V_2) .$$

Since the applied voltage $V = V_1 + V_2$, we obtain the following relationship between V and V_1:

$$V = V_1 + \psi_2 - \alpha \left[\frac{kT}{q} \left(e^{q(\psi_1 - V_1)/kT} - 1 \right) - (\psi_1 - V_1) \right], \tag{3.2}$$

where $\alpha \equiv \epsilon_1 N_{D1} / \epsilon_2 N_{D2}$.

From the point of view of understanding the physics of the isotype heterojunction, we are interested in establishing the voltage distributions $V_1(V)$ and $V_2(V)$ for a given applied voltage V. However, Equation 3.2 provides us only with the inverse relationship in the case of V_1, that is, knowing V_1 determines V, and attempting to obtain the direct dependence $V_1(V)$ from Equation 3.2 we are faced with a transcendental equation in V_1. Not only is this an issue in terms of the analytical solutions for the voltage distributions $V_1(V)$ and $V_2(V)$, but it also affects the derivation of the current density $J(V)$. When the underlying model for current transport is formulated in terms of V_1 and V_2 directly, via a particular mapping $J(V_1, V_2)$, as opposed to the special case $J(V_1, V_2) = J(V_1 + V_2) = J(V)$, it is necessary to explicitly provide $V_1(V)$ and $V_2(V)$ in order to arrive at a closed-form analytical expression for $J(V)$. This is in fact the situation we encounter in the derivation for the current density in an isotype heterojunction.

As done in Anderson's derivation and subsequent works (Sze and Ng, 2007; Chang, 1965; Anderson, 1962), the dominant current transport mechanism is assumed to be thermionic emission, neglecting quantum mechanical transmission and reflection, diffusion across the barrier, conduction through the bulk, recombination, interface states, as well as minority carrier transport. Thus, the current density is assumed to be of the

form

$$J(V) = J_0 e^{-q\psi_2/kT}\left[e^{qV_2/kT} - e^{-qV_1/kT}\right] = J_0 e^{-q\psi_2/kT} e^{-qV_1/kT}\left[e^{qV/kT} - 1\right], \qquad (3.3)$$

where $J_0 \equiv qN_{D2}\sqrt{kT/2\pi m_2^*}$ and m_2^* is the effective mass of the semiconductor in the region $l > 0$. Because of the explicit dependence of the current transport model on V_1 and V_2 here, it is necessary to circumvent the limitation imposed by the transcendental form of Equation 3.2. However, note that while Equation 3.2 is transcendental in V_1 for general α, it ceases to be so in the special case $\alpha = 1$, where it becomes easily analytically solvable without the need of any special functions. This is precisely why the works (Sze and Ng, 2007; Anderson, 1962; Chang, 1965) only consider that case. As has been previously mentioned, this constitutes a major limitation of the ensuing analytical solution, so it is desirable to find an alternate way to solve the problem. Fortunately, the underlying form of Equation 3.2 is part of a class of transcendental equations that can be solved with the Lambert W function, introduced in Chapter 2.

Note that Equation 3.2 can be rewritten in the form of Equation 2.7 with an appropriate choice of parameters, hence the solution to Equation 3.2 can be derived in a similar manner as Equation 2.8. More precisely, Equation 3.2 can be transformed into

$$(c - x)e^{c-x} = -be^c \qquad (3.4)$$

where $x = q(\psi_1 - V_1)/kT$, $b = \alpha/(\alpha - 1)$, and $c = [q(V - \psi)/kT(\alpha - 1)] - b$, and the case $\alpha \neq 1$ is treated separately. Applying $W(z)$ to both sides of Equation 3.4, we

obtain $x = c - W(-be^c)$. Reverting to the original variables and parameters this yields

$$V_1(V) = \psi_1 + \frac{kT}{q}\left\{W_j\left[-\frac{\alpha}{\alpha-1}\exp\left(\frac{1}{\alpha-1}\left[\frac{q(V-\psi)}{kT}-\alpha\right]\right)\right] - \frac{1}{\alpha-1}\left(\frac{q(V-\psi)}{kT}-\alpha\right)\right\}, \tag{3.5}$$

where at this point the solutions are still indexed in terms of the branches of $W(z)$, $W_j(z)$. In the case $\alpha = 1$ the relevant equation ceases to be transcendental and we can easily get

$$V_1(V;\alpha=1) = \psi_1 + \frac{kT}{q}\ln\left[\frac{kT/q}{(\psi + (kT/q) - V)}\right]. \tag{3.6}$$

3.1.1 Existence and Uniqueness of the General Solution

The derivation of Equation 3.5 gives the solution for $V_1(V)$ in terms of undetermined branches of the Lambert W function, leaving the question of existence and uniqueness of real solutions to be addressed, which is rigorously done in this Section. The arguments that follow are based on the properties of the Lambert W function introduced in Chapter 2.

The fact that for $\alpha = 1$ the solution in Equation 3.6 is expressed without the Lambert W function already suggests that the mathematical behaviour of the general solution is best understood by considering the cases $0 < \alpha < 1$ and $\alpha > 1$ separately. The relevant aspect to analyze in Equation 3.5 is the behaviour of the term $-be^c$ appearing as the argument of $W_j(z)$ (with the definitions of b and c as in Equation 3.4).

For the case $0 < \alpha < 1$, $b < 0$, so that the positivity of $-be^c$ leads to a *unique real solution* in terms of $W_0(x)$.

Consider now the case $\alpha > 1$. It is useful to write $c = d - b$ with $d \equiv q(V - \psi)/kT(\alpha - 1)$. Note that tacitly the physical model assumes $V < \psi$ (see also Equation 1.1), so that $d < 0$ and $e^d < 1$. But then, $e^c = e^d e^{-b} < e^{-b}$. Also, for $\alpha > 1$, $b > 1$, hence $-be^c > -be^{-b}$. Lastly, note that the function $f(x) = xe^x$ has a global minimum at the point $(-1, -1/e)$. Therefore, we have $0 > -be^c > -be^{-b} > -1/e$, which is precisely the domain where the two real branches of W, W_0 and W_{-1}, are simultaneously defined. In other words, the case $\alpha > 1$ contains *two mathematically possible real solutions* for $V_1(V)$ in terms of W_0 and W_{-1}.

While the mathematical arguments prove existence in this case, we must turn to the underlying physics to establish uniqueness. Choosing W_0 leads to an unbounded solution as $\alpha \to 1^+$, since $c \to -\infty$ and $-be^c \to 0^-$, leading to $[W_0(-be^c) - c] \to 0 - (-\infty)$. On the other hand, choosing W_{-1} we have $\alpha \to 1^+ \Rightarrow [W_{-1}(-be^c) - c] \to -\infty + \infty$, whose indeterminate form is not the most convenient for finding the limit. Instead, noting that from the definition of $W(x)$ we have $e^{W(x)} = x/W(x)$, then

$$\exp[\,[W_{-1}(-be^c) - c] = \frac{-be^c}{W_{-1}(-be^c)} e^{-c} = -\frac{b}{W_{-1}(-be^c)} \,.$$

As discussed in Chapter 2, for negative values of x near 0, $W_{-1}(x)$ can be approximated by an asymptotic expansion, the first term of which is $\ln(-x)$ (Corless *et al.*, 1996). Thus,

$$\begin{aligned}
\lim_{\alpha\to 1^+}\left[-\frac{b}{W_{-1}(-be^c)}\right]^{-1} &= -\lim_{\alpha\to 1^+}\frac{\ln b + c}{b} = -\left[\lim_{b\to\infty}\frac{\ln b}{b} + \lim_{\alpha\to 1^+}\frac{d-b}{b}\right] \\
&= -\lim_{\alpha\to 1^+}\frac{d}{b} + 1 = 1 - \lim_{\alpha\to 1^+}\frac{q(V-\psi)}{kT(\alpha-1)}\frac{\alpha-1}{\alpha} \\
&= 1 - \frac{q(V-\psi)}{kT} \,.
\end{aligned}$$

Hence,

$$\lim_{\alpha\to 1^+}\left[W_{-1}(-be^c)-c\right]=\ln\left(1-\frac{q(V-\psi)}{kT}\right)^{-1}=\ln\left(\frac{kT}{kT+q(\psi-V)}\right).$$

Recalling that $V_1=\psi_1+kT/q\,[W(-be^c)-c]$, the W_{-1} branch solution leads to

$$\lim_{\alpha\to 1^+}V_1=\psi_1+\frac{kT}{q}\ln\left(\frac{kT}{kT+q(\psi-V)}\right),$$

which is the solution obtained in Equation 3.6 for $\alpha=1$. Therefore, while W_0 leads to an unbounded solution as $\alpha\to 1^+$, the W_{-1} solution leads to continuity for $\alpha\in[1,\infty)$, providing us with a *unique, physically sensible solution for* $\alpha>1$.

3.1.2 The Complete Generalized Solution

The above observations allows us to construct the complete solution for all positive values of α as follows:

$$V_1(V)=\psi_1+\frac{kT}{q}\gamma(V;\alpha),\quad V_2(V)=V-V_1(V),$$
$$\gamma(V;\alpha)=\begin{cases}W_0(-be^c)-c & 0<\alpha<1\\ \ln\left[\dfrac{kT/q}{(\psi+(kT/q)-V)}\right] & \alpha=1\\ W_{-1}(-be^c)-c & \alpha>1\end{cases}\tag{3.7}$$

with b and c defined as before.

We can see that the function $\gamma(V;\alpha)$ is continuous in α on the entire domain $(0,\infty)$ from the following observations. First of all, each of the W branches is continuous on

its domain. Moreover, the arguments made in Section 3.1.1 already established that γ is right-continuous at $\alpha = 1$. To see left-continuity consider $W_0(-be^c) - c$ as $\alpha \to 1^-$. In this case, $c \to \infty$, $b \to -\infty$, and so $[W_0(-be^c) - c] \to W_0(\infty) - \infty = \infty - \infty$. The indeterminacy here can again be resolved by arguing in the same fashion as we did for the case of $\alpha \to 1^+$, while replacing $W_0(x)$ with $\ln(x)$ instead of $\ln(-x)$ in the limit $x \to \infty$, i.e.,

$$\begin{aligned}
\lim_{\alpha\to 1^-} \exp[W_0(-be^c) - c] &= \lim_{\alpha\to 1^-} -\frac{b}{W_0(-be^c)} = -\lim_{\alpha\to 1^-} \left[\frac{W_0(-be^c)}{b}\right]^{-1} \\
&= -\left[\lim_{\alpha\to 1^-} \frac{\ln(-b)}{b} + \frac{c}{b}\right]^{-1} = -\left[\lim_{b\to -\infty} \frac{\ln(-b)}{b} + \lim_{\alpha\to 1^-} \frac{d-b}{b}\right]^{-1} \\
&= -\left[-1 + \lim_{\alpha\to 1^-} \frac{q(V-\psi)}{kT(\alpha-1)} \frac{\alpha-1}{\alpha}\right]^{-1} \\
&= \left[1 - \frac{q(V-\psi)}{kT}\right]^{-1}
\end{aligned}$$

Taking logarithms and simplifying we have

$$\lim_{\alpha\to 1^-} V_1 = \psi_1 + \frac{kT}{q} \ln\left(\frac{kT}{kT + q(\psi - V)}\right),$$

establishing the left-continuity of $\gamma(V;\alpha)$ at $\alpha = 1$.

Note also that while the approach taken in the derivation of Equation 3.7 was to solve for $V_1(V)$ first and obtain $V_2 = V - V_1(V)$, choosing to isolate $V_2(V)$ would still require the use of the Lambert W function in solving for $V_1(V_2)$.

If we now employ the expression for $J(V)$ given in Equation 3.3, then the generalized current density becomes

$$J_G(V;\alpha) = J_0 e^{-q\psi/kT - \gamma(V;\alpha)} \left[e^{qV/kT} - 1\right]. \tag{3.8}$$

If we set $\alpha = 1$ and apply the condition $\psi \gg kT/q$, then Equation 3.8 reduces to Equation 1.1,

$$J_1(V) = J_0 \frac{q\psi}{kT} e^{-q\psi/kT} \left(1 - \frac{V}{\psi}\right) \left[e^{qV/kT} - 1\right] ,$$

i.e, the standard expression in the literature derived under the assumptions $\alpha = 1$ and $\psi \gg kT/q$ (e.g. Sze and Ng (2007); Chang (1965)).

A somewhat curious property of both Equation 3.8 and Equation 1.1 is that despite Equation 3.3 indicating the transport is dominated by ψ_2, the final current density expressions appear to depend only on $\psi = \psi_1 + \psi_2$. Note however that determining any of the pairs $(\alpha, \psi_1), (\alpha, \psi_2)$ or (α, ψ), completely determines the others. This is a consequence of the inherent relationship between ψ_1, ψ_2 and α that arises from considering Equation 3.2 under the condition $V = 0$, i.e.,

$$\psi_2(\psi_1, \alpha) = \alpha \left[\frac{kT}{q}\left(e^{q\psi_1/kT} - 1\right) - \psi_1\right] . \tag{3.9}$$

Equation 3.9 provides ψ_2 explicitly in terms of ψ_1, and the relationship can be inverted with the Lambert W function to obtain ψ_1 in terms of ψ_2:

$$\psi_1(\psi_2, \alpha) = -\frac{kT}{q}\left[W_{-1}\left(-e^{-y(\psi_2)}\right) + y(\psi_2)\right] , \tag{3.10}$$

where $y(\psi_2) \equiv (q\psi_2/\alpha kT) + 1$.

The branch choice in Equation 3.10 is again determined by appealing to the physical sensibility of the solution. Consider the case where ψ_2 is fixed and we allow α to vary in $(0, \infty)$. As $\alpha \to 0^+$, $y \to \infty$ and $-e^{-y(\psi_2)} \to 0^-$. On the other hand, as $\alpha \to \infty$,

$y \to 1$. Thus, the domain $(0,\infty)$ in α corresponds to $(-1/e, 0)$ in the argument of W in Equation 3.10, $-e^{-y(\psi_2)}$, which ensures the expression stays real, but also *a priori* not unique. However, a solution involving the branch W_0 would be unphysical. Recall that $W_0(x) > -1$ for all $x \in (-1/e, \infty)$, hence $W_0\left(-e^{-y(\psi_2)}\right) + y(\psi_2) > -1 + y(\psi_2) > 0$, since by definition $y > 1$, leading to negative values of ψ_1 for any $\alpha > 0$. On the other hand, $y > 1 \Rightarrow -ye^{-y} < -e^{-y} \Rightarrow -y > W_{-1}(-e^{-y})$, since W_{-1} is a decreasing function, which leads to the positivity of ψ_1 for all values of α. Thus, *there is a unique, physically relevant solution expressing ψ_1 in terms of ψ_2, given by the W_{-1} branch.*

Using $\psi = \psi_1 + \psi_2$ together with Equation 3.9 or Equation 3.10, we can rewrite Equation 3.8 in terms of either (α, ψ_1) or (α, ψ_2), respectively, with each particular form illuminating different physical aspects of the solution.

3.2 Contours of Constant Dominant Barrier ψ_2

3.2.1 Semiconductor-Metal Schottky Contacts as a Special Limit

Consider the set of contours of Equation 3.8 where ψ_2 is a fixed quantity and ψ is expressed as $\psi_1(\psi_2, \alpha) + \psi_2$. Examining the high α behaviour of the built-in potential, $\psi = \psi_1 + \psi_2$, as well as the current density, $J(V;\alpha)$, leads to an unexpected but physically very sensible result. As noted after Equation 3.10, $\alpha \to \infty \Rightarrow -e^{-y(\psi_2)} \to -1/e$, hence $\psi_1(\psi_2) \to -kT/q[W_{-1}(-1/e) + 1] = 0$ for any value of ψ_2. Also, considering Equation 3.7, note that as $\alpha \to \infty$, $b \to 1$ and $c \to -1$. Then we have $\alpha \to \infty \Rightarrow \gamma(V;\alpha) = \left[W_{-1}(-be^c) - c\right] \to W_{-1}(-1/e) - (-1) = 0$. Thus, for constant ψ_2, we

obtain,

$$\lim_{\alpha\to\infty} J_G(V;\alpha) = J_0\, e^{-q\psi_2/kT}\left[e^{qV/kT}-1\right]. \tag{3.11}$$

In a physical context, the definition $\alpha \equiv \epsilon_1 N_{D1}/\epsilon_2 N_{D2}$ entails a bound on the values of α, and Equation 3.11 is interpreted as an asymptotic behaviour. In the context where the limit $\alpha \gg 1$ is approached via $N_{D1} \gg N_{D2}$ for a fixed value of N_{D2}, this leads to a very physically intuitive result. Letting ϕ_B be the barrier height in the semiconductor SC_2 with reference to the Fermi-level, so that $q\phi_B = q\psi_2 + E_C - E_F$, then

$$\exp(-q\phi_B/kT) = \exp(-q\psi_2/kT)\exp[-(E_C-E_F)/kT] = \exp(-q\psi_2/kT)\frac{n}{N_C},$$

where n is the electron density, which can be replaced by N_{D2} in the semiconductor region SC_2. Thus, we can write Equation 3.11 as

$$\lim_{\alpha\to\infty} J_G(V;\alpha) = A^* T^2 e^{-q\phi_B/kT}\left[e^{qV/kT}-1\right], \tag{3.12}$$

where $A^* = 4\pi q m_2^* k^2/h^3$ is the effective Richardson's constant (Sze and Ng, 2007), and h is Planck's constant. Equation 3.12 is exactly the expression for the current density across a Schottky metal-semiconductor contact (Sze and Ng, 2007). This result is not surprising if we consider the associated band diagram. With ψ_1 going to zero while keeping ψ_2 a constant quantity, such a band diagram is indistinguishable from that of a metal-semiconductor Schottky contact.

However, we should not interpret Equation 3.12 in the context of the special case where the two materials SC_1 and SC_2 are fixed and the doping is altered. The reason is

that given a fixed set of materials SC_1 and SC_2, variations in α lead to variations in ψ_2, due to the fixed difference in their electron affinities. The limit is better interpreted as being reached through a path in the space of all isotype heterointerfaces via the particular contour of fixed ψ_2, allowing us to include the description of metal-semiconductor Schottky contacts as an asymptotic behaviour of a sequence of isotype heterojunctions.

3.2.2 Some Numerical Examples

Figure 3.1 shows the normalized current density J_G/J_0 as a function of V for a fixed $\psi_2 = 0.3$V, $T = 300$K, and different $\alpha-$contours, which all have a diode-like behaviour. Having fixed ψ_2, the absence of a strong variation for lower values of V further illustrates that the transport is mainly determined by this barrier, but it is also interesting to see that at higher voltages two asymptotic regimes in α start to be clearly identified, with the crossover near $\alpha = 10$.

Figure 3.2 explores the behaviour of $J_G(V;\alpha)$ normalized by the high α limit in Equation 3.11, and Figure 3.3 considers the ratio V_1/V, both as functions of α for fixed $\psi_2 = 0.3$V, $T = 300$K, and different values of V. Notice that lower α values lead to larger values of V_1, which decreases the current, as expected from Equation 3.3.

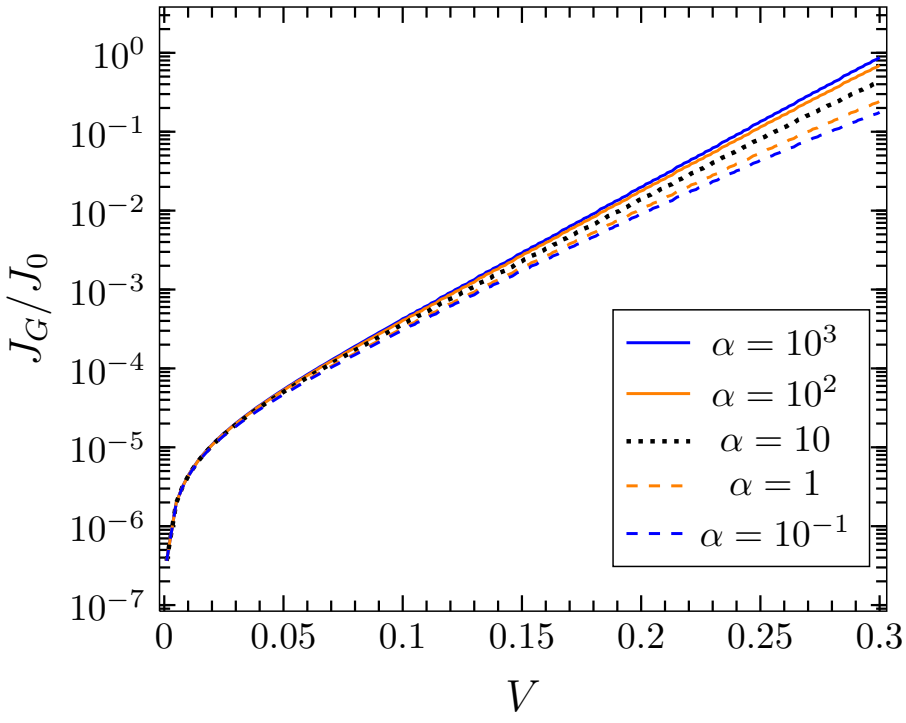

Figure 3.1: Normalized current density J_G/J_0 as a function of V for several values of the doping ratio α with fixed $\psi_2 = 0.3$V and $T = 300$K. Two asymptotic regimes can be identified for low and high values of α, with the crossover near $\alpha = 10$.

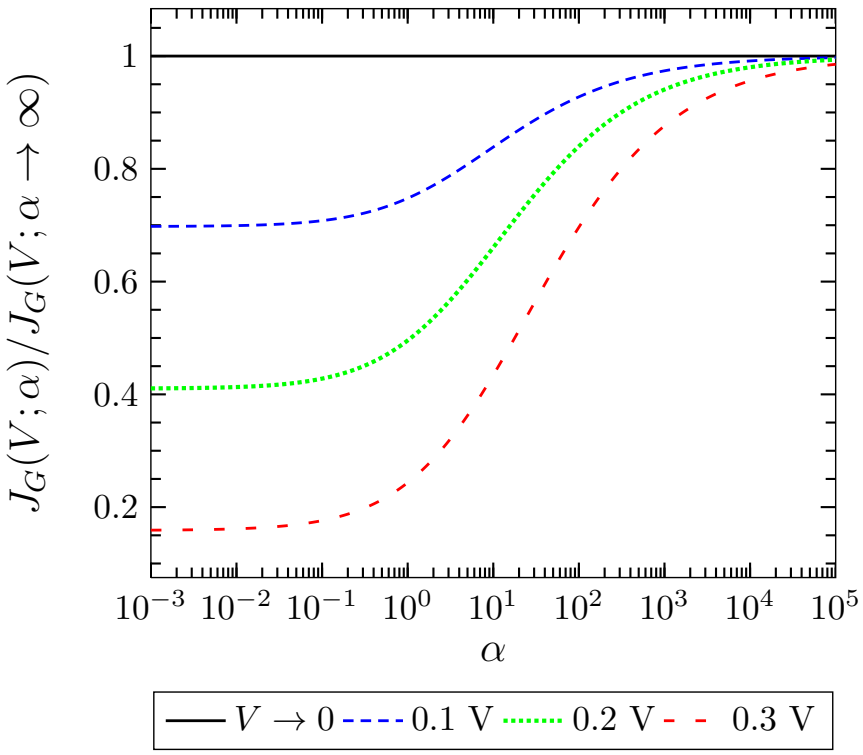

Figure 3.2: Generalized current density $J_G(V;\alpha)$ normalized by the high α Schottky limit, as a function of α for several values of the applied voltage V, $\psi_2 = 0.3$V and $T = 300$K.

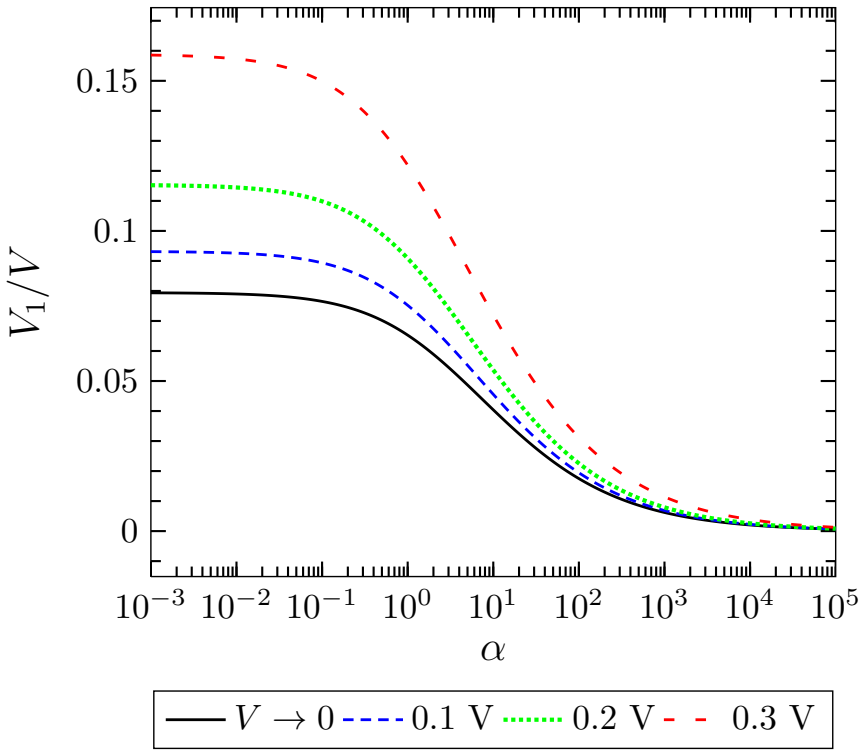

Figure 3.3: The voltage drop V_1 in SC_1 normalized by the applied voltage V, as a function of α for several values of the applied voltage V, $\psi_2 = 0.3$V and $T = 300$K.

3.2.3 Approximate Analytical Expression for Low α and fixed ψ_2

In light of the high α results in Section 3.2.1, it is also interesting to consider the low α behaviour of the barrier ψ_1 and the generalized current density $J(V;\alpha)$ under a fixed ψ_2. Consider again the expression for ψ_1 as a function of ψ_2 given in Equation 3.10,

$$\psi_1(\psi_2,\alpha) = -\frac{kT}{q}\left[W_{-1}\left(-e^{-y(\psi_2)}\right) + y(\psi_2)\right],$$

where $y(\psi_2) \equiv (q\psi_2/\alpha kT)+1$. As $\alpha \to 0^+$, the argument of W_{-1} goes to 0^-. Recall that $W_{-1}(x)$ can be approximated for negative x near zero with the asymptotic expression

$$W_{-1} = \ln(-x) - \ln(-\ln(-x)) .$$

Thus, near $\alpha \approx 0$, we can approximate

$$\begin{aligned}\psi_1(\psi_2) &= -\frac{kT}{q}\left[-y(\psi_2) - \ln(y(\psi_2)) + y(\psi_2)\right] = \frac{kT}{q}\ln\left(\frac{q\psi_2}{\alpha kT}+1\right)\\ &\approx \frac{kT}{q}\ln\left(\frac{q\psi_2}{\alpha kT}\right) .\end{aligned} \tag{3.13}$$

Figure 3.4 shows a plot of $\psi_1(\alpha)$ for $\psi_2 = 0.3$V. Note the low α logarithmic behaviour expected from Equation 3.13, as well as the fact that $\alpha \to \infty \Rightarrow \psi_1 \to 0$ discussed earlier.

Examining the low α behaviour of $\gamma(V;\alpha)$ in Equation 3.7, we have

$$\begin{aligned}\gamma(V;\alpha<1) &= W_0\left[-\frac{\alpha}{\alpha-1}\exp\left(\frac{1}{\alpha-1}\left[\frac{q(V-\psi)}{kT}-\alpha\right]\right)\right] - \frac{1}{\alpha-1}\left(\frac{q(V-\psi)}{kT}-\alpha\right)\\ &\approx W_0\left[\alpha\exp\left(\frac{q(\psi-V)}{kT}\right)\right] - \frac{q(\psi-V)}{kT} .\end{aligned}$$

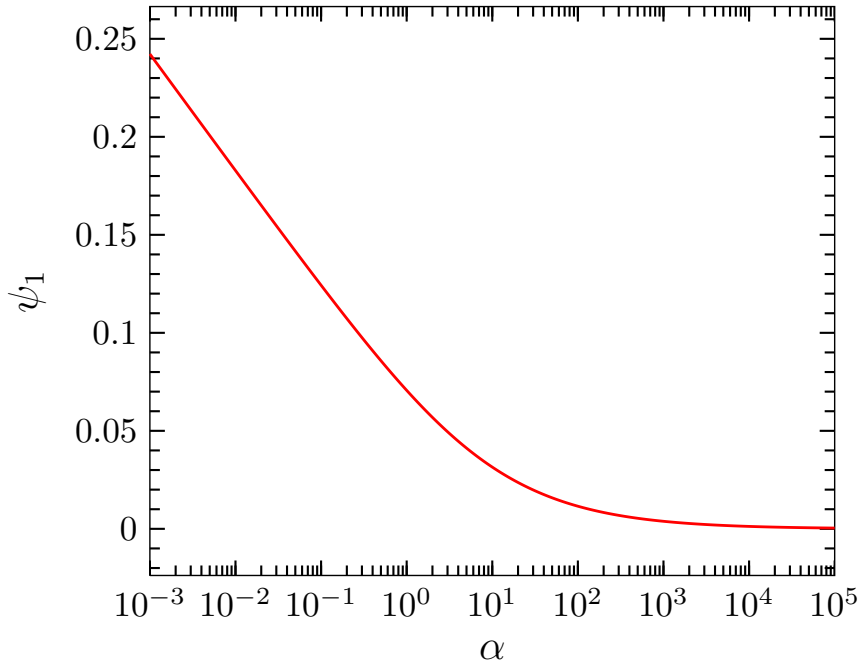

Figure 3.4: $\psi_1(\alpha)$ for $\psi_2 = 0.3$V.

But now

$$\frac{q(\psi - V)}{kT} \approx \frac{q}{kT}\left(\psi_2 + \frac{kT}{q}\ln\left(\frac{q\psi_2}{\alpha kT}\right) - V\right),$$

leading to

$$\begin{aligned}\gamma(V;\alpha \approx 0) &\approx W_0\left[\alpha \exp\left(\frac{q(\psi_2 - V)}{kT}\right)\frac{q\psi_2}{\alpha kT}\right] - \frac{q(\psi - V)}{kT} \\ &= W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2 - V)}{kT}\right)\right] - \frac{q(\psi - V)}{kT}. \end{aligned} \tag{3.14}$$

Applying this approximation to the generalized expression in Equation 3.8,

$$J_G(V;\alpha) = J_0 e^{-q\psi/kT - \gamma(V;\alpha)}\left[e^{qV/kT} - 1\right],$$

yields

$$J_G(V;\alpha\approx 0)\approx \exp\left(-W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2-V)}{kT}\right)\right]\right)\sqrt{\frac{kT}{2\pi m_2^*}}qN_{D2}\left[1-e^{-qV/kT}\right]. \tag{3.15}$$

Recall that the current density for the high-α Schottky limit is

$$J(\alpha\gg 1)=\sqrt{\frac{kT}{2\pi m_2^*}}qN_{D2}e^{-q\psi/kT}\left[e^{qV/kT}-1\right]\ .$$

Since it has already been shown rigorously that for fixed ψ_2, ψ_1 goes to zero in the high α limit, ψ can then be replaced with ψ_2. Then, for fixed ψ_2, the ratio of the high and low α asymptotic expressions for the current density becomes

$$\begin{aligned}\frac{J(\alpha\gg 1)}{J_G(\alpha\ll 1)}&\approx\frac{e^{-q\psi_2/kT}\left[e^{qV/kT}-1\right]}{\exp\left(-W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2-V)}{kT}\right)\right]\right)\left[1-e^{-qV/kT}\right]}\\&=\frac{e^{q/kT(V-\psi_2)}}{\exp\left(-W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2-V)}{kT}\right)\right]\right)}\end{aligned} \tag{3.16}$$

Consider now the low and high V limits of the ratio in Equation 3.16. In the limit $V\to 0$, we have

$$W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2-V)}{kT}\right)\right]\to\frac{q(\psi_2)}{kT},$$

since W is the inverse of the function $x\to xe^x$. Hence,

$$\lim_{V\to 0}\frac{J(\alpha\gg 1)}{J_G(\alpha\ll 1)}=\frac{e^{-q\psi_2/kT}}{e^{-q\psi_2/kT}}=1\,.$$

Thus, as $V \to 0$ the high and low α asymptotic expressions converge to zero at the same rate, as evidenced by the low V behaviour observed in Figure 3.1. In fact, Figure 3.1 also shows that below about 0.05V $\approx 2\ kT/q$ (at 300K), the current density is essentially constant in α.

At the other end, for higher voltages, $V \to \psi_2^-$ (recall that the physical model assumes $V < \psi$), we have

$$\frac{J(\alpha \gg 1)}{J_G(\alpha \ll 1)} = \exp\left(W_0\left(q\psi_2/kT\right)\right)\ , \tag{3.17}$$

which is greater than 1 by the positivity of $W_0(x)$ for $x > 0$. Thus, the closer the applied voltage is to ψ_2, the the more significant the increase in current in the high α regime, as evidenced by Figure 3.1 for the particular case $\psi_2 = 0.3$V. Note also that since W_0 is an increasing function, higher values of ψ_2 lead to a higher ratio in Equation 3.17.

As seen in Figure 3.3, V_1/V also exhibits asymptotic behaviours at both the high and low α regimes, increasing with decreasing α and increasing applied voltage V. It is therefore of interest to investigate the behaviour of V_1 for low α as we vary ψ_2 while keeping the applied voltage $V = \psi_2$. Figure 3.5 plots the expression for $V_1(\psi_2)$ obtained by substituting Equation 3.10 into Equation 3.5 and setting $\alpha = 10^{-3}$ and $V = \psi_2$. Note that for all $\psi_2 < 1$V, $V_1 < 0.07 \approx 2.7\ kT/q$ (at 300K).

It is also interesting to derive a low α analytical expression for V_1. If we apply the low α approximation to $\gamma(V;\alpha)$ in Equation 3.14 to the $V_1(V)$ solution in Equation 3.5,

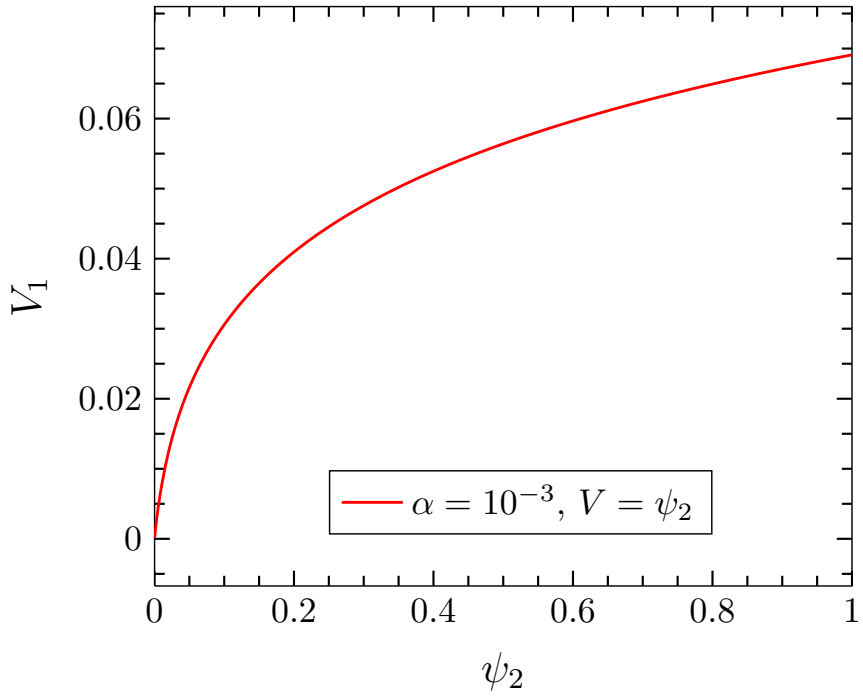

Figure 3.5: V_1 as function of ψ_2 (Equation 3.10 substituted into Equation 3.5) for fixed $\alpha = 10^{-3}$ and the applied voltage $V = \psi_2$.

we have

$$V_1 \approx \psi_1 + \frac{kT}{q}\left\{W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2 - V)}{kT}\right)\right] - \frac{q(\psi - V)}{kT}\right\}$$
$$= \frac{kT}{q}W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q(\psi_2 - V)}{kT}\right)\right] - (\psi_2 - V)\,. \tag{3.18}$$

Note that for $V \to 0$, Equation 3.18 yields

$$V_1 \to \frac{kT}{q}W_0\left[\frac{q\psi_2}{kT}\exp\left(\frac{q\psi_2}{kT}\right)\right] - \psi_2 = \psi_2 - \psi_2 = 0\,,$$

which verifies the consistency of the expression. In the case $V \to \psi_2$, we obtain

$$V_1(\psi_2) = \frac{kT}{q} W_0\left(\frac{q\psi_2}{kT}\right) . \tag{3.19}$$

The accuracy of the low α approximation in Equation 3.18 is quite remarkable when compared to the non-approximate general result plotted in Figure 3.5 (Equation 3.10 substituted into Equation 3.5), which is a significantly more complex expression. Figure 3.6 plots the error between the approximation in Equation 3.19 and the complete expression plotted in Figure 3.5.

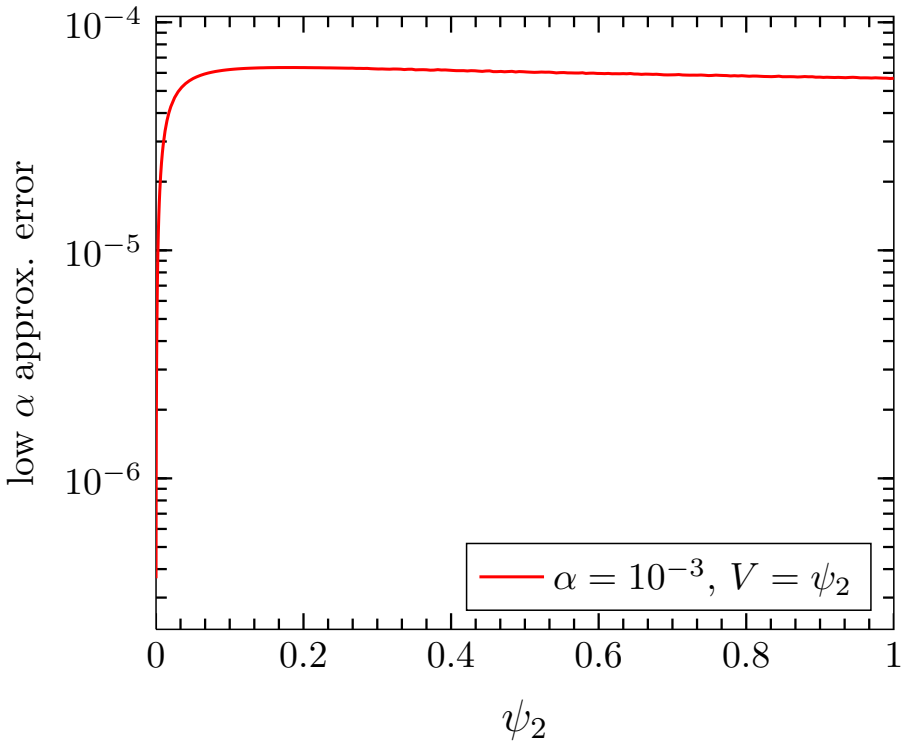

Figure 3.6: Error between the low α approximation for V_1 as a function of ψ_2, with the applied voltage $V = \psi_2$ (Equation 3.19), and the complete expression plotted in Figure 3.5 (Equation 3.10 substituted into Equation 3.5). Here the approximation very slightly overestimates the value of the complete expression.

3.3 A Simpler Approximate Analytical Solution for Low Applied Voltage Valid for all α

Note that for low values of V we can arrive at a simpler approximate analytical expression for $J(V;\alpha)$. Considering Equation 3.2, we can linearize the term $\exp(-qV_1/kT)$ to $1-qV_1/kT$, and rewrite

$$\begin{aligned} V &= V_1 + \psi_2 - \alpha\left[\frac{kT}{q}\left(e^{q\psi_1/kT}\left(1-qV_1/kT\right)-1\right)-(\psi_1-V_1)\right] \\ &= \left(1+\alpha e^{q\psi_1/kT}-\alpha\right)V_1 + \psi_2 - \alpha\left[\frac{kT}{q}\left(e^{q\psi_1/kT}-1\right)-\psi_1\right] \\ &= \left(1+\alpha e^{q\psi_1/kT}-\alpha\right)V_1\,, \end{aligned}$$

since the term independent of V_1 is identically zero by Equation 3.9. By rewriting the term $\alpha e^{q\psi_1/kT}-\alpha$ also using Equation 3.9, an equivalent form is

$$\frac{V_1}{V} = \left[1+\frac{q(\psi_2+\alpha\psi_1)}{kT}\right]^{-1},$$

which in turn leads to the current density

$$J(V;\alpha) = J_0 e^{-q\psi_2/kT}\left(e^{qV/kT}-1\right)\left(1-\frac{V}{kT/q+\psi_2+\alpha\psi_1}\right), \tag{3.20}$$

where the ψ_2 or ψ_1 dependency can be removed using Equation 3.9 or Equation 3.10, as previously discussed.

Equation 3.20 can be used to analytically estimate the crossover point, α^*, between the low and high α asymptotic regimes. Before developing that point, we need to

determine the behaviour of the limits $\alpha\psi_1(\alpha)$ as $\alpha \to 0, \infty$. From the low α approximation to $\psi_1(\psi_2)$ given in Equation 3.13, we have $\psi_1(\alpha) \sim \ln(1/\alpha)$, hence $\alpha \to 0 \Rightarrow \alpha\psi_1(\alpha) \to 0$. Figure 3.7 shows a plot of $\alpha\psi_1(\alpha)$ for $\psi_2 = 0.3$V, which also suggests that $\alpha \to \infty \Rightarrow \alpha\psi_1(\alpha) \to \infty$, which can be verified analytically. Recall that in Section 3.2.1 it was shown that $\alpha \to \infty \Rightarrow \psi_1(\alpha) \to 0$. Thus, for large α we may Taylor expand the term $\exp(q\psi_1/kT)$ in Equation 3.9 so that

$$\psi_2 = \alpha\left[\frac{kT}{q}\left(1 + \frac{q\psi_1}{kT} + \frac{1}{2}\left(\frac{q\psi_1}{kT}\right)^2 - 1\right) - \psi_1\right] = \frac{\alpha}{2}\frac{q\psi_1^2}{kT} \Leftrightarrow \psi_1 = \sqrt{\frac{2kT\psi_2}{q\alpha}}.$$

Thus, for large α, $\alpha\psi_1(\alpha) \sim \alpha\alpha^{-1/2} = \alpha^{1/2}$, hence $\alpha \to \infty \Rightarrow \alpha\psi_1(\alpha) \to \infty$.

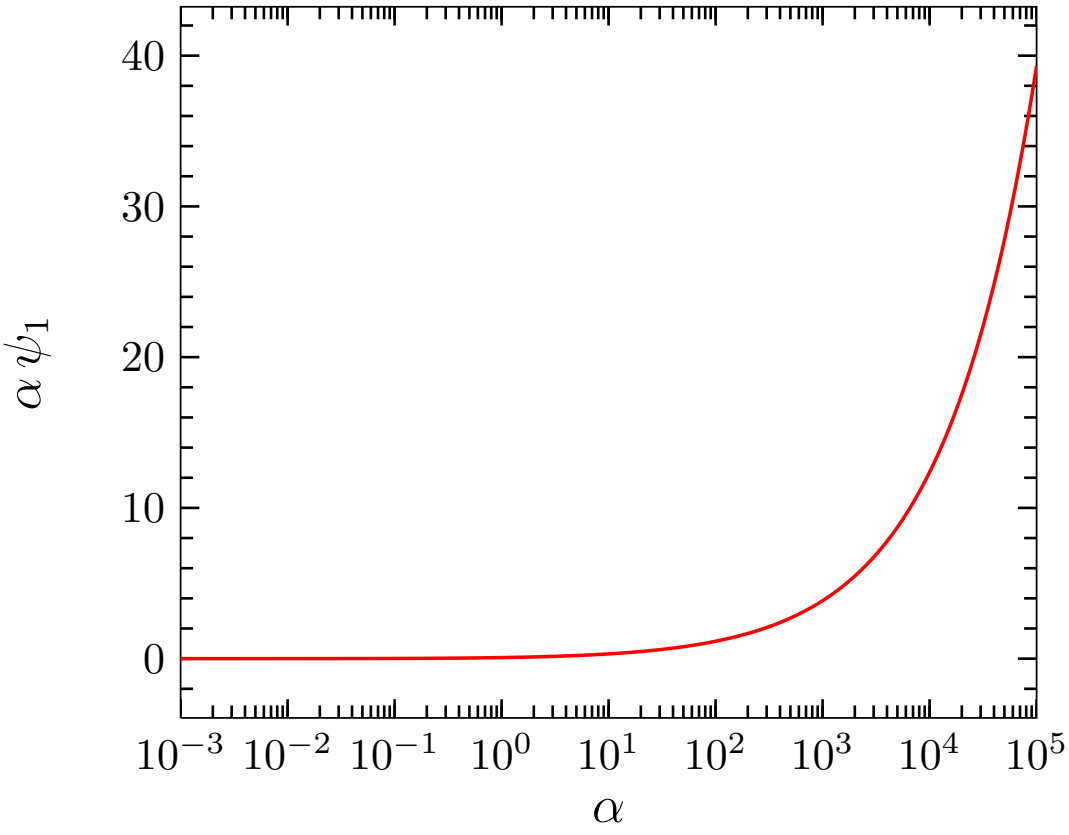

Figure 3.7: $\alpha\psi_1(\alpha)$ for $\psi_2 = 0.3$V.

Now that it has been established that for fixed ψ_2, $\alpha\psi_1 \to 0, \infty$ as $\alpha \to 0, \infty$, respectively, we can formulate an equation to estimate the crossover point, α^*, between the high α and the low α regimes. This is done by demanding that $J_G(V, \alpha^*)$ be the average of the high α and low α behaviour of Equation 3.20. Thus α^* must satisfy the equation

$$1 - \frac{V}{kT/q + \psi_2 + \alpha^*\psi_1} = \frac{1}{2}\left[1 - \frac{V}{kT/q + \psi_2} + 1\right] \Leftrightarrow \alpha^*\psi_1(\psi_2, \alpha^*) = \frac{kT}{q} + \psi_2 \,. \tag{3.21}$$

For $\psi_2 = 0.3$V and $T = 300$K, the solution is $\alpha^* \approx 10.6$, confirming the observation regarding the crossover in Figure 3.1 and Figure 3.2. Figure 3.8 shows the behaviour of α^* as a function of ψ_2, generated from the numerical solutions of Equation 3.21. It is also interesting that for all values of ψ_2, $\alpha^* > 2.58$, demonstrating that $\alpha = 1$ is never a crossover point. Moreover, as indicated by Figure 3.2 and Figure 3.3, the $\alpha = 1$ case is not fully in the asymptotic low α regime even for $\alpha^* = 10.6$. These observations suggest that the case $\alpha = 1$ is not really a physically illuminating or central case, and the assumption $\alpha = 1$ is only justifiable from the simplicity of the ensuing mathematics.

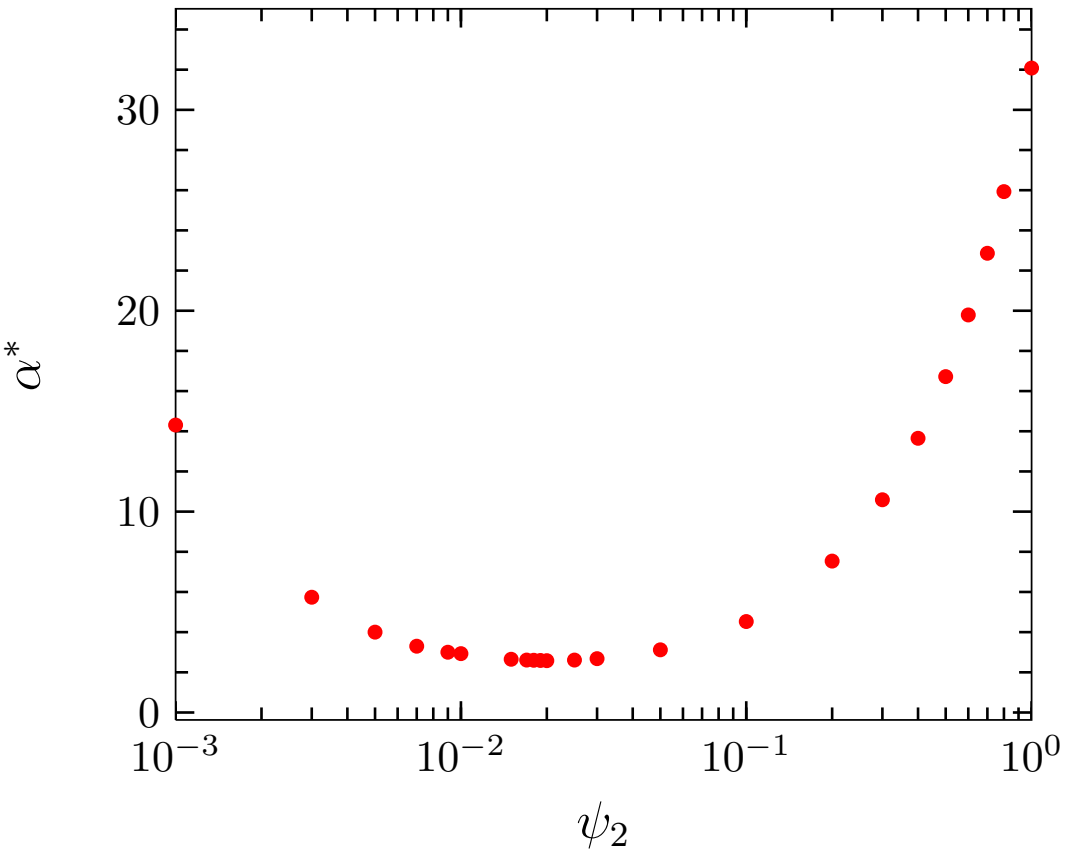

Figure 3.8: Values of the crossover point, α^*, as a function of ψ_2.

3.4 Other Relevant Contours

Besides the fixed ψ_2 contours considered in the previous sections, it is also of interest to examine the mathematical behaviour of $J_G(V;\alpha)$ for other contours. Figure 3.9 illustrates the behaviour of $J_G(V;\alpha)/J_0$ as a function of α when $\psi = 0.35$V. Note that $J_G(V,\alpha)$ has gone from being an *increasing* function of α under the fixed ψ_2 contour illustrated in Figure 3.1, where $\psi = \psi_1(\psi_2) + \psi_2$, to a *decreasing* function of α in the fixed ψ contour of Figure 3.9. In addition, for a fixed pair of α values, the separation between different curves decreases with increases in V, which is also the *opposite* of what was is observed in Figure 3.1. Figure 3.10 shows a plot of the ratio of $J_G(V,\alpha)$ to $J_G(V,\alpha\to\infty)$, in analogy with Figure 3.2, but now on the fixed $\psi = 0.35$V contour.

The change in the dependence of $J_G(V;\alpha)$ on α can be explained by the behaviour of the barriers under this contour. When fixing ψ, ψ_1 and ψ_2 are implicitly varying to accommodate the variations in α while keeping the net $\psi = \psi_1 + \psi_2 = 0.35$V. This can be seen in Figure 3.11, which is generated by solving for ψ_1 in the equation $\psi_2(\psi_1) + \psi_1 = \psi$, using Equation 3.9, and then obtaining $\psi_2 = \psi - \psi_1$. The fact that the dominant barrier ψ_2 is now a monotonically increasing function of α explains the behaviour observed above for $J_G(V;\alpha)$.

Another natural question to consider is whether we can "decompose" the dependence on the barriers ψ_1 and ψ_2 into functions of the underlying material parameters and the doping levels, and explore more specific contours that way. For example, one could fix one of the semiconductors and vary the material properties of the other to arrive at an optimized interface. This would be very useful from the point of view of applications, since one could compile readily available data for the expected current densities for a given combination of materials, removing the need to know the

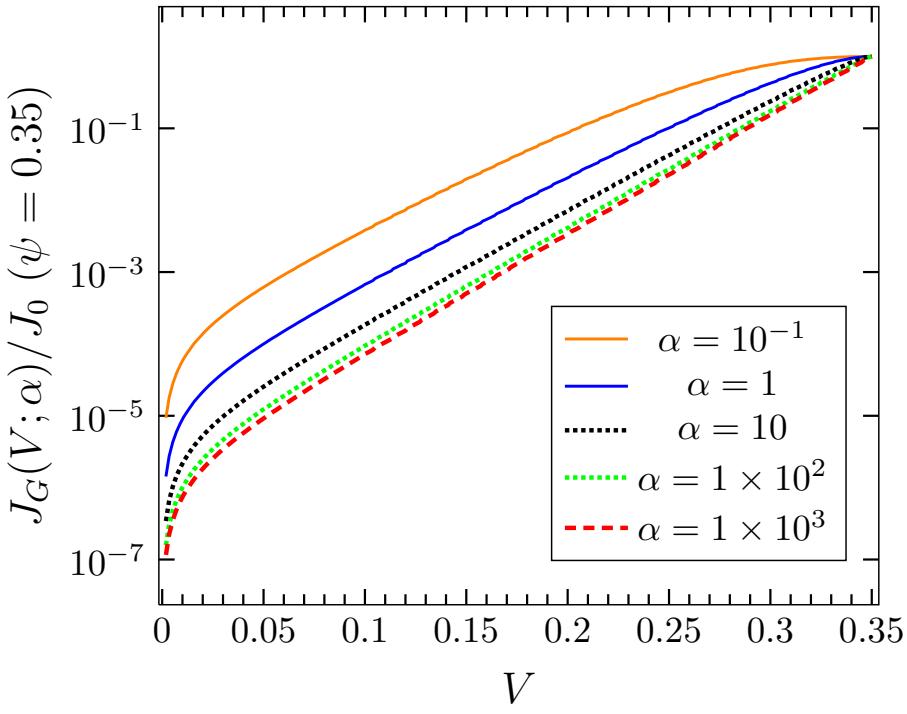

Figure 3.9: $J_G(V;\alpha)/J_0$ as a function of V for fixed values of α, with $T = 300$K and $\psi = \psi_1 + \psi_2 = 0.35$V.

quantitative details of each band diagram.

Unfortunately, the information about the band structure of each semiconductor is not sufficient (Franciosi and de Walle, 1996). Considering an individual semiconductor, SC, the band structure can be calculated in relation to a reference potential, which is usually the average of the electrostatic potential in SC, $\overline{V}$. The positions of the valence and conduction bands can be measured as being a certain distance away from $\overline{V}$. Because of the reference to an underlying average potential, when we consider two different semiconductors, it becomes necessary to be able to determine the reference potential in each of the materials as measured by an absolute energy scale, otherwise any given potential will not be comparable across the two semiconductors. However, there is no absolute reference for the average potential in an infinite solid. This is

a consequence of the long range nature of the Coloumb interaction, as explained in Franciosi and de Walle (1996).

Nevertheless, the relative potential of two specific materials can be determined experimentally, as in the classic case of the Ge-GaAs heterojunction, and used to determine the ψ values.

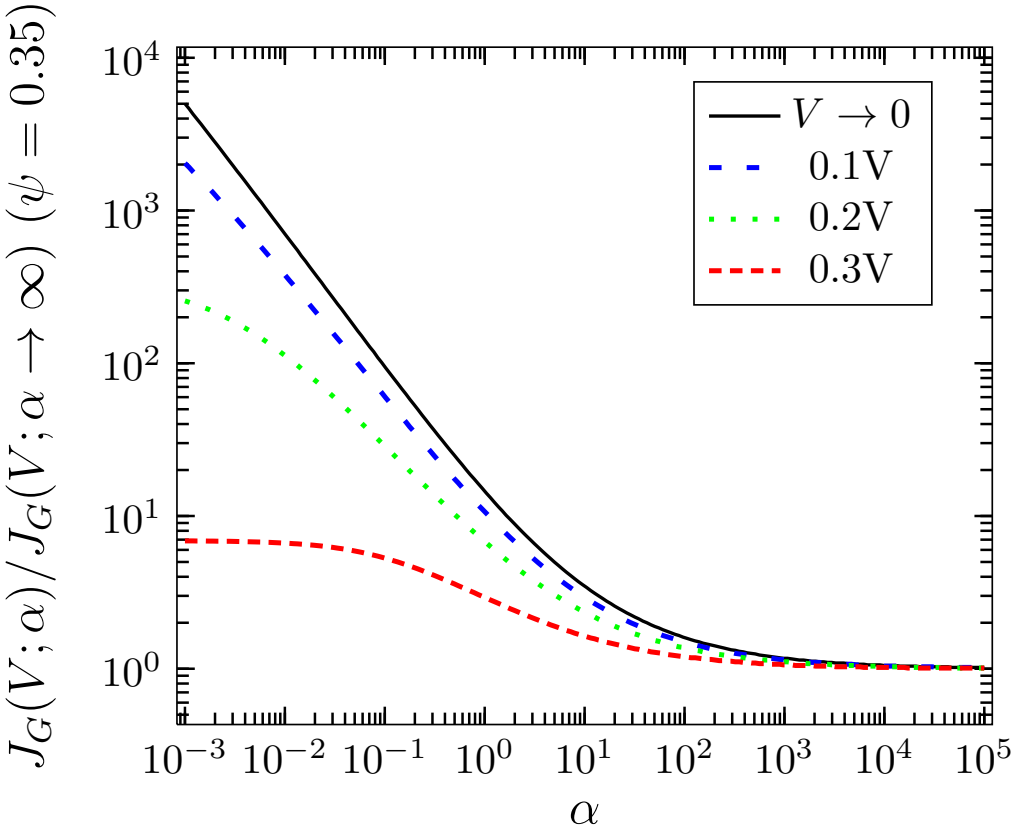

Figure 3.10: $J_G(V;\alpha)/J_0$ as a function of α for fixed values of V, with $\psi=\psi_1+\psi_2=$ 0.3V, and $T=300$K.

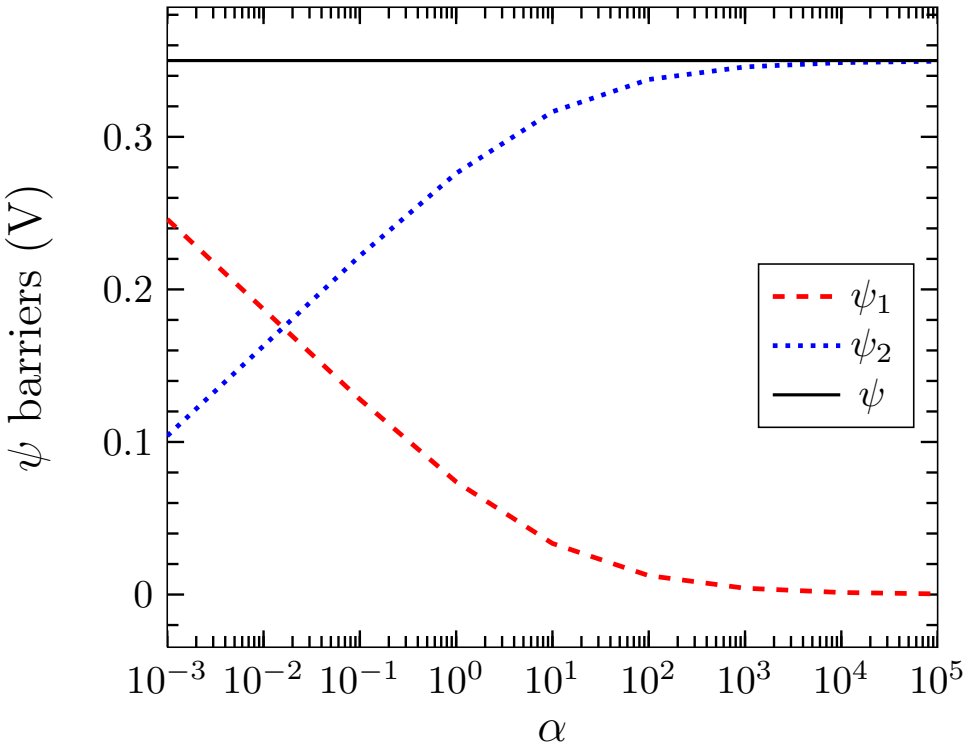

Figure 3.11: Behaviour of ψ_1 and ψ_2 as functions of α for fixed values of $\psi = \psi_1 + \psi_2$.

Chapter 4

An Application to Solar Cells

As discussed in Section 1.1, isotype heterojunctions play an important role in the design of multijunction solar cells. Thus, understanding the resistance across these junctions becomes a crucial step in the optimization of the entire device.

Referring back to Figure 1.2, we can model the device as the following series connection: top cell–isotype heterojunction–tunnel junction–bottom cell. The particular place of the tunnel junction and the isotype heterojunction (as well as its type) will depend on which particular silicon multijunction cell design is implemented. However, in all cases it is desirable that the voltage drop across both the isotype heterojunction and the tunnel junction be as small as possible. While the analysis of the tunnel junction is outside the scope of this thesis, we can use the results obtained here for isotype heterojunctions to identify parameter spaces that meet acceptable design guidelines.

In the context of multijunction solar cells, it is of particular interest to understand the resistance under low applied voltage, which can be approximated by the differential resistance at zero applied voltage. We can therefore make use of the low voltage analytical expression for $J(V)$ in Equation 3.20 to derive the near zero differential

resistance. This leads to

$$R_0 \equiv \left[\left(\frac{\partial J_G(V;\alpha)}{\partial V} \right)_{V=0} \right]^{-1} = \frac{h}{q^2 N_{D2}} \left(\frac{N_{C2}}{2} \right)^{1/3} e^{q\psi_2/kT} . \tag{4.1}$$

Thus, for low V, the differential resistance has a straightforward exponential dependence on the dominant barrier ψ_2, and it is independent of α.

For a typical triple junction solar cell, at 1 sun, one can estimate the short-circuit current $J_{sc} \sim 15\text{mA/cm}^2$ and the open-circuit voltage $V_{oc} \sim 2.5V$ (King *et al.*, 2007), which results in a resistance of $166.7\,\Omega\text{cm}^2$. The presence of the isotype heterojunction resistance will lower the operating voltage to $V_{oc} - I_{sc}R_0$. If we require that the voltage drop as a result of the resistance across the isotype heterojunction be no more than 1% of the cell's V_{oc}, then we need $R_0 < 1.67\Omega\ \text{cm}^2$. However, because the tunnel junction will also provide additional resistance, a more conservative number may be $1\Omega\text{cm}^2$. At 1000 suns, $J_{sc} \sim 15A/\text{cm}^2$ and $V_{oc} \sim 3.0$V, so that the isotype resistance must be kept below $2 \times 10^{-3}\,\Omega\ \text{cm}^2$, but again, the number must be reduced to account for extra resistance from the tunnel junction, so $1 \times 10^{-3}\,\Omega\ \text{cm}^2$ is a better bound.

Figure 4.1 displays a plot of R_0 as a function of ψ_2 for $N_{D2} = 1 \times 10^{16}\text{cm}^{-3}$, $T = 300$K, and various values of N_{C2}. The '1 sun' and '1000 suns' lines correspond to the discussion above. We can see that for typical values of N_{C2}, the threshold value for the 1 sun requirement is $\psi_2 \approx 0.3$V, whereas for much higher concentrations even a ψ_2 as low as 0.2V may not satisfy the 1% efficiency drop bound. In order to obtain a more precise bound one needs to know or estimate the resistance contribution from the tunnel junction, which is not considered in this work.

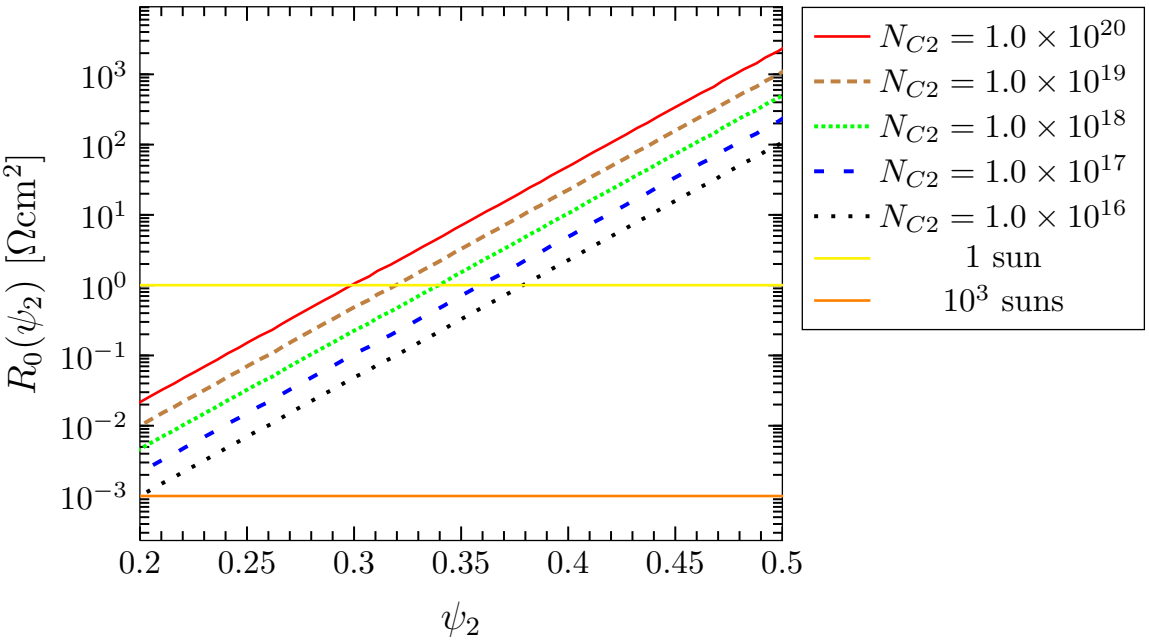

Figure 4.1: Differential Resistance at $V = 0$ as a function of ψ_2 for $N_{D2} = 1 \times 10^{16}$cm^{-3}, $T = 300$K, and various values of N_{C2} (in cm^{-3}).

Chapter 5

Conclusion

This thesis has presented an analytical study of the I-V characteristics of isotype heterojunctions. The main contribution of this work is the derivation of a generalized analytical expression for the current density $J(V;\alpha)$ across an isotype heterojunction for arbitrary doping concentrations parametrized via $\alpha \equiv \epsilon_1 N_{D1}/\epsilon_2 N_{D2}$. This result constitutes a significant expansion to the expression originally introduced by Anderson (Anderson, 1962), which relies on the assumption that the doping concentration ratio $\alpha = 1$. Such an assumption has been maintained even in recent well-known semiconductor device physics references such as Sze and Ng (2007). The generalization presented here makes use of the recently rediscovered Lambert W function, which allows for the analytical solution of the boundary conditions associated with the electrostatics of isotype heterojunctions. It is worth noting that the voltage distribution determined by Equation 3.7 depends only on these heterojunction interface conditions, and not on the particulars of the current transport mechanism. Hence, the generalization can be readily carried out for any current density $J(V_1, V_2)$ derived under the assumptions

of Figure 1.3 and Equation 3.1. Also, it has been shown that the generalized expression introduced here mathematically contains Anderson's expression as the special case $\alpha = 1$.

Doping concentration is one of the crucial controllable design parameters in semiconductor devices. As such, the results in this work have important practical implications.

Moreover, the generalization presented here is also of theoretical significance. It has been shown that the expression for the current density across a metal-semiconductor Schottky contact also follows from the generalized current density for isotype heterojunctions in the limit $\alpha \to \infty$ under a fixed barrier height ψ_2, which is consistent with the expected physics of this case. This demonstrates that metal-semiconductor Schottky contacts and isotype heterojunctions are special cases of a general heterointerface, whose I-V characteristic is now analytically modelled in an unified and mathematically consistent fashion. This finding is specially significant considering that isotype semiconductor heterojunctions and metal-semiconductor Schottky contacts have been traditionally described independently, e.g., (Sze and Ng, 2007).

Lastly, as shown in Chapter 4, the analysis developed here leads to a readily usable expression for the differential resistance of a general isotype heterojunction near zero voltage, which is an important device optimization parameter when considering multijunction solar cell designs that implement isotype heterojunctions.

Future Work

There are a few avenues to explore to further develop the ideas presented in this work, as briefly suggested below.

Additional transport terms in the generalized current density

As pointed out above, the derivation of the voltage distribution Equation 3.7 depends only on the assumptions of Figure 1.3 and Equation 3.1. Having determined the voltage drops across the interface, V_1 and V_2, for all applied voltages V, one can easily add additional current transport terms to generate a more complete generalized expression. This could be especially useful in applications where the band diagram has a barrier that is neither too low nor too high for either diffusion or thermionic emission becoming the dominant transport mechanism, or if one wishes to include additional terms related to quantum mechanical tunneling and reflection, as for example in a high barrier that starts to become thin enough for these to matter. Regardless of the particulars of the transport mechanism, so long as assumptions Figure 1.3 and Equation 3.1 are met, adding any transport function $J_T(V_1, V_2)$ will result in a generalized $J(V)$ from the solutions $V_1(V)$ and $V_2(V)$ obtained in this work.

Coupling the generalized current density with an empirical characterization of the barrier for a particular set of materials

Recall the discussion in Section 3.4 about the impossibility of predicting the barrier ψ in the isotype heterojunction from only the knowledge of each material's band parameters. However, considering a specific set of materials, it may be possible to have empirical characterizations that formulate a mathematical relationship between the barrier ψ and the doping concentration on each semiconductor. For materials of the form $A_xB_{1-x}C$, it may even be possible to obtain ψ also as a function of x. Such relationships can then be coupled with the analytical results of this work to arrive at a function $J(V)$, parametrized by variables that can in principle be optimized for a

particular design goal.

Adding Fermi Dirac corrections

Since one of the assumptions in the derivation of the generalized current density is that carriers follow Maxwell-Boltzmann statistics in the region SC_1 (Figure 1.3), it may be of interest to attempt the solution of $V_1(V)$ and $V_2(V)$ from the electrostatic conditions at the isotype heterojunction with additional correction terms associated with Fermi-Dirac statistics. One possibility is to express the relevant Fermi integrals as power series (e.g. through the polylogarithm function), and try to carry out the inversion via the Lagrange inversion theorem that led to the series derivation of the Lambert W function, as discussed in Corless *et al.* (1996).

Bibliography

Anderson, R. (1962). Experiments on Ge-GaAs heterojunctions. *Solid-State Electronics*, **5**(5), 341 – 351.

Caillol, J. M. (2003). Some applications of the Lambert W function to classical statistical mechanics. *Journal of Physics A: Mathematical and General*, **36**, 10431–10442.

Casey, H. C. and Panish, M. B. (1978). *Heterostructure lasers: Fundamental principles - Part A*. Academic Press, New York.

Chandrashekar, R. and Segar, J. (2013). Adiabatic thermostatistics of the two parameter entropy and the role of Lambert's -function in its applications. *Physica A: Statistical Mechanics and its Applications*, **392**(19), 4299 – 4315.

Chang, L. (1965). The conduction properties of Ge-GaAs1-xPx n-n heterojunctions. *Solid-State Electronics*, **8**(9), 721 – 728.

Corless, R. M., Gonnet, G. H., Hare, D. E. G., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W function. *Advances in Computational Mathematics*, **5**(1), 329–359.

European Photovoltaic Industry Association (EPIA) (2013). Global market outlook for photovoltaics 2013-2017. http://www.epia.org/home/.

Farrugia, P. S., Mann, R. B., and Scott, T. C. (2007). N -body gravity and the Schrödinger equation. *Classical and Quantum Gravity*, **24**(18), 4647.

Fonash, S. (2010). *Solar Cell Device Physics*. Academic Press.

Franciosi, A. and de Walle, C. G. V. (1996). Heterojunction band offset engineering. *Surface Science Reports*, **25**(1-4), 1–140.

Germani, C. and Liguori, M. (2009). Matching WMAP 3-year results with the cosmological slingshot primordial spectrum. *General Relativity and Gravitation*, **41**(1), 191–201.

Jain, A. and Kapoor, A. (2005a). A new approach to study organic solar cell using Lambert W-function. *Solar Energy Materials and Solar Cells*, **86**(2), 197 – 205.

Jain, A. and Kapoor, A. (2005b). A new method to determine the diode ideality factor of real solar cell using Lambert W-function. *Solar Energy Materials and Solar Cells*, **85**(3), 391 – 396.

Jain, A., Sharma, S., and Kapoor, A. (2006). Solar cell array parameters using Lambert W-function. *Solar Energy Materials and Solar Cells*, **90**(1), 25 – 31.

Jenn, D. (2002). Applications of the Lambert W function in electromagnetics. *Antennas and Propagation Magazine, IEEE*, **44**(3), 139–142.

Jizba, P. and Arimitsu, T. (2006). Towards Information Theory for q-Nonextensive Statistics Without q-Deformed Distributions. *Physica A*, **365**, 76–84.

Kast, S. M. and Tomazic, D. (2012). Communication: An exact bound on the bridge function in integral equation theories. *Journal of Chemical Physics*, **137**(17), 171102.

King, R. R., Law, D. C., Edmondson, K. M., Fetzer, C. M., Kinsey, G. S., Yoon, H., Sherif, R. A., and Karam, N. H. (2007). 40% efficient metamorphic GaInP/GaInAs/Ge multijunction solar cells. *Applied Physics Letters*, **90**(18), 183516.

Leite, M. S., Woo, R. L., Munday, J. N., Hong, W. D., Mesropian, S., Law, D. C., and Atwater, H. A. (2013). Towards an optimized all lattice-matched InAlAs/InGaAsP/InGaAs multijunction solar cell with efficiency $> 50\%$. *Applied Physics Letters*, **102**(3), 033901.

Lutz, E. (2005). Analytical results for a Fokker–Planck equation in the small noise limit. *American Journal of Physics*, **73**(10), 968–972.

Mann, R. B. and Ohta, T. (1997). Exact solution for the metric and the motion of two bodies in (1+1)-Dimensional gravity. *Physical Review D*, **55**, 4723–4747.

Milnes, A. G. and Feucht, D. L. (1972). *Heterojunctions and metal-semiconductor junctions*. Academic Press, New York.

Molli, M., Venkataramaniah, K., and Valluri, S. R. (2011). The polylogarithm and the Lambert W functions in thermoelectrics. *Canadian Journal of Physics*, **89**(11), 1171–1178.

Morales, D. A. (2011). A generalization on projectile motion with linear resistance. *Canadian Journal of Physics*, **89**(12), 1233–1250.

Nelson, J. (2003). *Physics of Solar Cells*. Imperial College Press, London.

Sharp Corporation (2013). Sharp develops concentrator solar cell with world's highest conversion efficiency of 44.4%. http://sharp-world.com/corporate/news/130614.html.

Starikov, E. B. (2010). Many faces of entropy or Bayesian statistical mechanics. *ChemPhysChem*, **11**(16), 3387–3394.

Sze, S. M. and Ng, K. K. (2007). *Physics of Semiconductor Devices*. Wiley-Interscience, Hoboken, 3rd edition.

Tanguay, J., Gil, M., Jeffrey, D. J., and Valluri, S. R. (2010). D-dimensional bose gases and the Lambert W function. *Journal of Mathematical Physics*, **51**(12), 123303.

Valluri, S. R., Jeffrey, D. J., and Corless, R. M. (2000). Some Applications of the Lambert W Function to Physics. *Canadian Journal of Physics*, **78**, 823–831.

Valluri, S. R., Gil, M., Jeffrey, D. J., and Basu, S. (2009). The Lambert W function and quantum statistics. *Journal of Mathematical Physics*, **50**(10), 102103.

Veberic, D. (2010). Having fun with Lambert W(x) function. *CoRR*, **abs/1003.1628**.

Weisse, J. (2010). Concentrated solar photovoltaics. http://large.stanford.edu/courses/2010/ph240/weisse2/.

Williams, B. W. (2005). Exact solutions of a Schrödinger equation based on the Lambert function. *Physics Letters A*, **334**(2-3), 117 – 122.