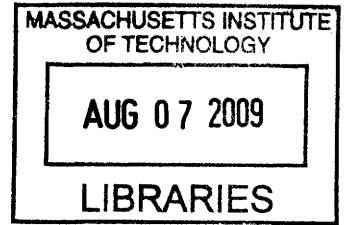


# Generalized Drift-Diffusion for Microscopic Thermoelectricity

by

Parthiban Santhanam



B.S. Engineering Physics, University of California at Berkeley (2006)

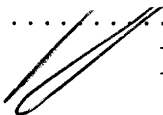
Submitted to the  
Department of Electrical Engineering and Computer Science  
in partial fulfillment of the requirements for the degree of  
Master of Science in Electrical Engineering and Computer Science  
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2009

© Massachusetts Institute of Technology 2009. All rights reserved.

Author .....  
Department of Electrical Engineering and Computer Science  
May 22, 2009

Certified by .....  
  
Rajeev J. Ram  
Professor  
Thesis Supervisor

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

**ARCHIVES**



# Generalized Drift-Diffusion for Microscopic Thermoelectricity

by

Parthiban Santhanam

Submitted to the Department of Electrical Engineering and Computer Science  
on May 22, 2009, in partial fulfillment of the  
requirements for the degree of  
Master of Science in Electrical Engineering and Computer Science

## Abstract

Although thermoelectric elements increasingly incorporate nano-scale features in similar material systems as other micro-electronic devices, the former are described in the language of irreversible thermodynamics while devices such as heterojunction bipolar transistors and semiconductor lasers are often described with the drift-diffusion equations. We present a microscopic description of the thermoelectric effects using a generalization of the common drift-diffusion formulation of semi-classical transport. We then replicate these basic results in a commercial device simulation package to explore Peltier cooling at a basic  $p$ - $n$  junction. This framework should enable the design of spatially-inhomogeneous thermoelectric elements and internally-cooled micro-electronic devices.

Thesis Supervisor: Rajeev J. Ram  
Title: Professor



## Acknowledgments

I would like to begin by thanking Rajeev Ram, who has served as my research advisor for the past two years. My interactions with him have shaped my approach to many things, from problems in applied science to developing relationships with labmates.

I would also like to take this opportunity to specifically mention the tremendous intellectual freedom that I've been afforded, and as anyone who's shared the lab with me knows, not at the expense of advisor face time. Such free access to a person like Rajeev has been a gift. The freedom to choose coursework and research based primarily on interest (thanks to DoD-NDSEG as well for that) has combined with the seriously diverse range of active research projects in the group to create a lot of learning and a lot of fun.

And of course that fun usually comes in the form of my labmates, all of whom I'm quite fond of by this point. I could go down the list and thank people for discussions and their humor, but in the interest of not giving people license to embarrass me in their theses, I'll save it for my doctoral thesis.

I would also like to thank my close friends and roommates, who have made MIT a part of my life I'm sure to miss. They have helped with everything from quantum field theory to learning to accept my receding hairline. I know Shawn will get uppity if I don't mention him specifically, so thanks go out to him for his genuine friendship.

Finally, I would like to thank my family. To Amma, I'd like to say that I know sometimes I get busy out here, so I'm sorry that I don't keep in touch better. To Akka, I'm consistently surprised at how we seem to be getting closer every year. Even if it's getting old to hear it again, I am deeply satisfied to have such a close relationship with my sister. To Appa, it's hard to thank someone you've modeled so much of yourself after, but I'll try anyway- thanks.

Oh yeah, one last thing- thanks, Amma and Appa, *foreverything*.

# Contents

<b>1</b>	<b>Introduction</b>	<b>21</b>
1.1	Thermoelectric Phenomena . . . . .	22
1.2	Thermoelectric Devices . . . . .	25
1.2.1	Bulk Material Analysis . . . . .	25
1.2.2	Thermodynamic Onsager Analysis . . . . .	29
1.2.3	Microscopic Derivation of Transport Coefficients . . . . .	31
1.3	Thermoelectric Applications . . . . .	37
1.3.1	Direct Power-Generation . . . . .	38
1.3.2	Macro-Scale Waste Heat Recovery . . . . .	40
1.3.3	Micro-scale Wireless Sensor Power . . . . .	41
1.3.4	Cooling and Temperature-Stabilization . . . . .	42
1.4	Thermal Self-Management Applications . . . . .	42
1.5	Thesis Outline . . . . .	43
<b>2</b>	<b>Reciprocal-Space Picture of Thermoelectric Transport</b>	<b>45</b>
2.1	Introduction and Overview . . . . .	45
2.2	The New Choice of Variables . . . . .	46
2.3	The Relaxation-Time Approximation . . . . .	47
2.4	Drift and Diffusion in the Boltzmann Limit . . . . .	55
2.4.1	Field Driven Transport . . . . .	55
2.4.2	Density Gradient Driven Transport . . . . .	57
2.4.3	Temperature Gradient Driven Transport . . . . .	58
2.5	Outside the Boltzmann Limit . . . . .	62
2.6	Summary and Conclusions . . . . .	67

<b>3</b>	<b>Real-Space Picture of Thermoelectric Transport</b>	<b>69</b>
3.1	Introduction and Overview . . . . .	69
3.2	The Microscopic Seebeck Effect . . . . .	71
3.2.1	The Balance Equations . . . . .	72
3.2.2	Self-Consistent Solution . . . . .	77
3.2.3	The Various Relaxation Times . . . . .	83
3.3	Seebeck with Energy-Dependent Relaxation Time . . . . .	89
3.4	Decomposing the Seebeck Effect . . . . .	93
3.5	More Realistic Material Systems . . . . .	96
3.5.1	Optimal Doping for Power-Generation . . . . .	96
3.5.2	A Note on Complex Bandstructures . . . . .	97
3.6	The Microscopic Peltier Effect . . . . .	99
3.6.1	The Energy Balance Equation and Peltier Heat . . . . .	100
3.6.2	Homojunctions . . . . .	103
3.7	Summary and Conclusions . . . . .	106
<b>4</b>	<b>Applying the Framework</b>	<b>109</b>
4.1	Chapter Overview . . . . .	109
4.2	Application Space for the Theory . . . . .	109
4.3	Transition to Commercial Device Simulators . . . . .	111
4.3.1	Replicating the Seebeck Effect . . . . .	112
4.3.2	Replicating the Peltier Effect . . . . .	116
4.3.3	Translation of Sentaurus Variables . . . . .	118
4.4	Internal Cooling at a $p$ - $n$ Junction . . . . .	120
4.5	Maximal Internal Cooling in a Short $p$ - $n$ Diode . . . . .	124
4.6	Summary and Conclusions . . . . .	127
<b>5</b>	<b>Summary and Future Work</b>	<b>129</b>
5.1	Summary and Chapter Overview . . . . .	129
5.2	Future Experimental Work . . . . .	130
5.2.1	Experimental Verification of Internal Cooling Model . . . . .	130

5.2.2	Experimental Proposals . . . . .	133
5.3	Future Theoretical Work . . . . .	135
<b>A</b>	<b>The Semi-Classical Picture</b>	<b>139</b>
A.1	The Boltzmann-Transport Equation . . . . .	139
A.2	The BBGKY Heirarchy . . . . .	141
<b>B</b>	<b>Non-Equilibrium Thermodynamics of Thermoelectricity</b>	<b>143</b>
B.1	Continuous Systems . . . . .	144
B.2	The Onsager Reciprocity . . . . .	147
B.3	The Onsager Relations . . . . .	148
B.4	The Kelvin Relations . . . . .	149
<b>C</b>	<b>Simulation Source Code</b>	<b>151</b>
C.1	Numerical Solver for Microscopic Transport Coefficients Outside the Boltzmann Limit (MATLAB) . . . . .	151
C.2	Seebeck Generalized Drift-Diffusion Solver (MATLAB) . . . . .	162
C.3	Peltier Generalized Drift-Diffusion Solver (MATLAB) . . . . .	178
	<b>Bibliography</b>	<b>191</b>



# List of Figures

Figure 1-1 : Left: Phenomenological depiction of the Seebeck effect. Right: Plot of open-circuit voltage with temperature difference across a commercial thermoelectric generator consisting of 142 elements connected electrically in series and thermally in parallel. This measurement was made in a demonstration power-generator built by the author. Note that the observed “average Seebeck” above is diminished from the material property value because of thermal parasitics. . . . .	23
Figure 1-2 : Left: Phenomenological depiction of the Peltier effect. Right: Plot of steady-state $\Delta T$ across a commercial thermoelectric cooler versus current $I$ required to maintain it. This Micro-Pelt MPC-D303 has 8 elements electrically in series and thermally in parallel. Plot taken from ( <a href="http://www.micropelt.com/products/peltier_cooler.php">http://www.micropelt.com/products/peltier_cooler.php</a> ). Note that while the Peltier effect is defined under isothermal conditions, driving a current through a TE material as in the figure at left will result in heat transport from the left contact to the right contact (for positive $\Pi$ ), giving rise to a temperature difference $\Delta T$ . In steady-state, this rate of heat transport exactly cancels the rate of ordinary heat conduction giving rise to a maximum $\Delta T$ as shown in the plot at right and derived in Section 1.2.1. . . . .	24
Figure 1-3 : Room-temperature ZT achievements with time [1]. Text in the figure should read “Discovery of compounds with $ZT > 1$ .” . . .	37
Figure 1-4 : One of three RTGs used to power the Cassini Orbiter, launched 1997 [17]. . . . .	38

Figure 1-5 : Demonstration thermoelectric generator with large thermal parasitics, built by the author. . . . .	39
Figure 1-6 : Operating temperature profile for exhaust system of a BMW passenger vehicle [21]. . . . .	40
Figure 1-7 : Proposed site for waste-heat recovery in an automotive TEG system. Image from ( <a href="http://www.bsst.com/automotive-waste-heat-recovery.php">http://www.bsst.com/automotive-waste-heat-recovery.php</a> ). . . . .	41
Figure 1-8 : Consideration of internal Peltier cooling at the base-emitter junction of heterojunction bipolar transistors (HBTs) could lead to improved device designs [2]. . . . .	42
Figure 1-9 : By examining the Peltier heating and cooling present in typical diode lasers, whose band diagram and Peltier heat generation profile shown left, Pipe proposed a design [2] for the ICICLE (Injection Current Internally Cooled Light Emitter), depicted at right. . . . .	43
Figure 2-1 : Plot of phase-space density as a function of $\vec{E}$ -directed wave-vector for a distribution of negatively charged carriers under a positive electric field. The average wave-vector for the deviated distribution is $\langle k \rangle = -4.56e7$ , indicating a flux of carriers to the left. Parameters for plot are as follows: $\tau = 1ps$ , $m^* = 0.063m_e$ , $T = 300K$ , and $n_D = 1e16 \text{ cm}^{-3}$ . The choice of ionized donor concentration puts the extrinsic semiconductor in the Boltzmann limit. . . . .	50
Figure 2-2 : Plot of phase-space density as a function of $\nabla n$ -directed wave-vector. The average wave-vector for the deviated distribution is $\langle k \rangle = +3.93e7$ , indicating a flux of carriers to the right. The situation is identical to that in Figure 2-1 except that $\vec{E} = 0$ and $\nabla n \neq 0$ . . . . .	51

Figure 2-3	: Plot of phase-space density as a function of $\nabla T$ -directed wave-vector. Note that although the peak is moved to the left, the average wave-vector for the deviated distribution is $\langle k \rangle = +1.57e7$ , indicating a flux of carriers to the right. The situation is identical to that in Figure 2-1 except that $\vec{E} = 0$ and $\nabla T \neq 0$ . . . . .	52
Figure 2-4	: Plot of $f_0$ arguments in (2.3) for $\tau(\vec{k}) = \tau_0$ . Positive-velocity phase-space volume elements come from negative relative locations and negative-velocity elements come from positive relative locations. The positions whose actual occupancy these origination-locations inform can be seen by following any branch to zero momentum where the mean-free-path vanishes. . . . .	53
Figure 2-5	: Plot of $f_0$ arguments in (2.3) for $\tau(\vec{k}) = \tau_{th} \left( \frac{ \vec{k} }{\sqrt{2}k_{th}} \right)^p$ , where $p$ is known as the scattering parameter. The general description of the $p = 0$ lines are the same as in Figure 2-4. For $\vec{k}$ along the plotted direction $\hat{z}$ , we see that the origination shift follows a simple power-law in $ \vec{k} $ . For the $\vec{k}_\perp \neq 0$ case, the behavior at small $k_z$ resembles that of the $p = 0$ case, while at large $k_z$ the origination shift follows the $\vec{k}_\perp = 0, p \neq 0$ result. . . . .	54
Figure 2-6	: Left: Conductivity as a function of carrier density. Right: Ratio of conductivity to the value expected from extrapolating the Boltzmann-limit expression. Note that the relaxation time $\tau$ is assumed here to be independent of $n_D$ . . . . .	64
Figure 2-7	: Left: Diffusivity as a function of carrier density. Right: Ratio of diffusivity to the value expected from extrapolating the Boltzmann-limit expression. Again, the relaxation time $\tau$ is assumed here to be independent of $n_D$ . . . . .	64
Figure 2-8	: Left: Soret coefficient as a function of carrier density. Right: Ratio of Soret coefficient to the value expected from extrapolating the Boltzmann-limit expression. Again, the relaxation time $\tau$ is assumed here to be independent of $n_D$ . . . . .	65

Figure 2-9 : $E_F$ versus $n_D$ for n-GaAs. Spherical parabolic bandstructure assumed, with $m^*$ taken as $0.063m_0$ . . . . .	65
Figure 3-1 : Visual depiction of Seebeck effect. Left: density-driven picture. Right: field-driven picture. . . . .	70
Figure 3-2 : The carrier density profile for Seebeck effect at $T_{\text{avg}} = 300\text{K}$ , shown with $\Delta T = 1\text{K}$ . There is a buildup of free charge carriers on the cold-side (right) and a depletion on the hot-side (left). Parameters for the simulation taken from n-GaAs doped $n_D = 1\text{e}16 \text{ cm}^{-3}$ : $m^* = 0.063m_e$ , $\epsilon_r = 10.6$ . For reasons explained in the text, we assume a fixed value for thermal conductivity $\kappa$ and take energy-independent momentum-relaxation time $\tau_m = 0.3\text{ps}$ . . . . .	78
Figure 3-3 : The source-specified currents sum to satisfy net flux balance at each point in space, as required by the steady-state electrical open-circuit boundary conditions in the Seebeck effect. All parameters for this simulation are identical to that from Figure 3-2. . . . .	79
Figure 3-4 : The band diagram for the Seebeck effect demonstrates that in the Boltzmann limit, only a fraction of the thermoelectric voltage can be ascribed to a gradient in the band-edge, and therefore gradient in the vacuum electric potential, through the bulk of the device. The $86 \mu\text{V}/\text{K}$ limit explained in Section 3.4 can be seen, as $\Delta T = 1\text{K}$ and $\Delta E_C \approx 86 \mu\text{V}$ . All parameters for this simulation are identical to that from Figure 3-2. . . . .	80

Figure 3-5 : Carrier density profiles for Seebeck effect at 300K with  $\Delta T = 1\text{K}$ . As the ionized dopant concentration is increased, the greater density of free carriers serves to screen the edge effects on shorter length-scales. Just as in the determination of the Debye length, this screening length is the result of competition between drift and diffusion. It is therefore intuitive that the characteristic screening length depends on carrier density in the same way as does the Debye length ( $\propto 1/\sqrt{n_D}$ ). . . . . 81

Figure 3-6 : The numerical calculations for the Seebeck coefficient done via the generalized drift-diffusion framework explained in this chapter agree with the analytical results from the differential-conductivity method of Section 1.2.3 across a wide range of doping concentrations and temperatures. The calculations in the above figure are for n-GaAs, and express agreement in the Boltzmann limit. We will discuss higher ionized dopant concentrations in Section 3.5.1. . . . . 81

Figure 3-7 : Qualitative diagram of electric potential as a function of space near the surface of a thermoelectric element under a small applied temperature gradient. All but  $86 \mu\text{V/K}$  of the potential drop should be found within a few Debye lengths of the contacts. . . . . 83

Figure 3-8 : Plot showing agreement across a range of doping densities between generalized drift-diffusion model and differential conductivity model for Seebeck coefficient as a function of scattering parameter  $r$ . The scattering parameter is defined as the exponent of energy-dependence in mean-free-path:  $\lambda(E) = \lambda_{th} \left( \frac{E}{k_B T} \right)^r$ . All calculations are for n-GaAs at 300K. . . . . 89

Figure 3-9 : The source-specified current density profile for the Seebeck effect in a material with scattering time  $\tau \propto |\vec{k}|^5$ , indicative of ionized-impurity scattering [24]. All other parameters are identical to Figure 3-3, including  $\tau_{th} = \tau(|\vec{k}| = \sqrt{m^* k_B T / \hbar})$ . . . . . 92

Figure 3-10 : The source-specified current density profile for the Seebeck effect in a material with scattering time $\tau \propto  \vec{k} ^{-1}$ , indicative of ionized-impurity scattering [24]. All other parameters are identical to Figure 3-3, including $\tau_{th} = \tau( \vec{k}  = \sqrt{m^*k_B T/\hbar})$ . . . . .	92
Figure 3-11 : The so-called thermoelectric power factor ( $\alpha^2\sigma$ ) which enters directly into $ZT$ as seen in (1.4), experiences a maximum as a function of doping near the concentration required for the equilibrium Fermi level to fall near the band-edge. These plots express the results of the computational generalized drift-diffusion solver, using results for the transport coefficients from Section 2.5. Left: Power factor versus ionized dopant density. Right: Power factor versus equilibrium $E_F(n_D)$ measured relative to $E_C$ . . . . .	96
Figure 3-12 : Material $ZT$ s with temperature for various technologically-relevant materials. Left: n-type. Right: p-type. Image taken from [20]. . . . .	98
Figure 3-13 : Band-edge profiles for $n - n^+ - n$ structure at 300K with currents of +1 kA/cm <sup>2</sup> (left) and -1 kA/cm <sup>2</sup> (right) flowing. The doping concentrations of both the $n$ ( $n_D = 1e15$ cm <sup>-3</sup> ) and $n^+$ ( $n_D = 4e15$ cm <sup>-3</sup> ) remain in the Boltzmann limit. Other material parameters are taken from GaAs: $m^* = 0.063m_e$ , $\epsilon_r = 10.6$ , $\tau_m = 0.3ps$ . The isothermal boundary condition of the Peltier effect is enforced, which is equivalent to $\kappa \rightarrow \infty$ . . . . .	104
Figure 3-14 : Rate of total heat density production, including both Peltier and Joule contributions, in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), some of the heat density production flips sign but some does not. . . . .	104
Figure 3-15 : Rate of Peltier-only heat density production in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), all of the Peltier heat density production flips sign. . . . .	105

Figure 3-16 : Rate of Joule-only heat density production in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), none of the Joule heat density production flips sign. 105

Figure 4-1 : Left: carrier density profile for the Seebeck effect simulation in Sentaurus. Right: carrier density profile for the Seebeck effect simulation in generalized drift-diffusion code in Section C.2. Just as in the simulation from Section 3.2, the left boundary is fixed at 300.5 K and the right boundary is fixed at 299.5 K, leading to a temperature gradient of 1 K/ $\mu\text{m}$ . The remaining parameters used in both simulations were also identical, and can be found in the caption of Figure 3-2. . . . . 113

Figure 4-2 : Electric field profile for the same Sentaurus Seebeck effect simulation as presented in Figure 4-1. We find that for energy-independent scattering times, the magnitude of the field away from the boundaries asymptotically reaches  $\approx 86 \mu\text{V}/\text{K} \times 1 \text{ K}/\mu\text{m} = 0.86 \text{ V}/\text{cm}$ . . . . . 114

Figure 4-3 : Carrier density profile resulting from invalid Sentaurus simulation of the Seebeck effect using the thermodynamic model. The ionized dopant concentration  $n_D = 1.5\text{e}18 \text{ cm}^{-3}$  corresponds to the doping level for peak power-factor, which is outside the Boltzmann limit. The software converges to a non-physical solution with free carriers piled up on the hot-side instead of the cold-side. All simulation parameters other than  $n_D$  are identical to those found in Figure 4-1. . . . . 114

Figure 4-4 : Results of driving a current of  $\pm 1 \text{ kA/cm}^2$  through an  $n\text{-}n^+\text{-}n$  GaAs homojunction at 300K in Sentaurus. All other parameters are taken from GaAs, and are the same as those used in Figure 3-13. Left:  $J > 0$ . Right:  $J < 0$ . In each case, the conduction band-edge profile is plotted with units shown on the left, along with the donor-atom and carrier density profiles with units shown on the right. 116

Figure 4-5 : Total heat density generation rate ( $\vec{J} \cdot \vec{E}$ ) profile, where cooling appears as negative. Plot is for the same simulation as Figure 4-4, which uses Sentaurus to model the  $\pm 1 \text{ kA/cm}^2$  current-bias of an  $n\text{-}n^+\text{-}n$  homojunction. . . . . 117

Figure 4-6 : Plot showing decomposition of total heat density generation rate into Peltier and Joule terms. Again, cooling appears as negative heat generation. The thick solid line indicates Peltier heat; the thin solid line indicates Joule heat; the dashed line represents total heat, equal to the sum of the two. Plot is for the same simulation as Figure 4-4, which uses Sentaurus to model the  $\pm 1 \text{ kA/cm}^2$  current-bias of an  $n\text{-}n^+\text{-}n$  homojunction. . . . . 117

Figure 4-7 : Depiction of Peltier heat exchange contributed by electrons and holes as a function of space. Since  $E_F$  for the two bands only re-equilibrate near the contacts, the picture takes the short-length approximation. Image from [2]. . . . . 121



Figure 4-8 : Profile of the heat density generation rate in a steady-state forward-biased GaAs  $p$ - $n$  diode. The solid line indicates the heat density generation rate with cooling shown as negative, and the corresponding scale is found on the right y-axis. The dashed lines indicate the profile of the conduction and valence band-edges measured relative to the equilibrium Fermi level at zero bias. The junction between  $p$ -doped region ( $n_A = 1e18$ ) and  $n$ -doped region ( $n_D = 1e18$ ) is modeled as abrupt, resulting in  $V_{\text{built-in}} \approx 1.49\text{V}$ . Diode as shown is forward biased at  $1\text{V}$ , resulting in a total current density of  $\approx 4.888 \text{ mA/cm}^2$ . Simulation taken with the length of the  $p$ - and  $n$ -doped regions as  $500\text{nm}$  each, though the region plotted above extends only  $100\text{nm}$  on each side of the junction. All material parameters for GaAs same as those taken in Figure 3-2. . . . . 123

Figure 4-9 : Diagrammatic description of heat generation rates in a forward-biased short-length  $p$ - $n$  diode. The area corresponding to  $IV$  (left block) represents the total Joule heat throughout the device and the area corresponding to  $I(V - V_{\text{built-in}})$  (right block) represents the Peltier heat which is transported from the junction to the contacts. 125

Figure 4-10 : Heat density generation rate profile for the same device as in Figure 4-8, but with a  $1.5\text{V}$  bias applied instead of  $1\text{V}$ , resulting in a current density of  $444.0 \text{ kA/cm}^2$  instead of  $4.888 \text{ mA/cm}^2$ . This plot is for the entire device, not just a small region surrounding the junction as in Figure 4-8. The solid line indicates magnitude of heat density generation rate, plotted on a log scale (right y-axis) so that cooling and heating both appear positively. The region near the junction experiences net cooling, but beyond the minima of heat generation rate magnitude the device experiences net heating. The dashed lines again indicate the band-edge profiles (left y-axis). . . . 126

Figure 4-11 : Plot of total heat generation rate, given by integrating the total heat density generation rate over the 1 $\mu\text{m}$ length of the $p$ - $n$ diode simulated in Figure 4-8, versus bias voltage. For sufficiently large bias, there is net heating. Note that this calculation omits heating at the contacts, permitting net cooling for $V_{\text{applied}}$ below $\approx 1.488\text{V}$ or $J_{\text{driven}}$ below $\approx 342 \text{ kA/cm}^2$ . . . . .	127
Figure 5-1 : Layer structure of InGaAs diode on which initial junction cooling experiments were done. Doping concentrations in $\text{cm}^{-3}$ . Image from [2]. . . . .	130
Figure 5-2 : Dependence of temperature on input power density ( $J_{\text{driven}} \times V_{\text{applied}}$ ) at junction (squares) and 2 $\mu\text{m}$ into the substrate (circles). Image courtesy of Dietrich Lürßen and Rajeev Ram. . . . .	131
Figure 5-3 : Band-edge profile for simulated structure under 0.6V, 1.055 $\text{kA/cm}^2$ bias. The left edge corresponds to the top contact and the right edge corresponds to a point 2 $\mu\text{m}$ into the substrate . . . . .	132
Figure 5-4 : Left: Sentaurus simulation results for temperature profile of InGaAs diode structure under 0.6V bias. Right: current density profile. This current density corresponds to $\approx 630 \text{ W/cm}^2$ . . . . .	132
Figure 5-5 : Proposed setup for an experiment to measure the scattering parameter governing electron dynamics in an $n$ -type thermoelectric material. The material would need to be doped weakly to remain in the Boltzmann regime, but the results may remain valid at the higher doping levels where $ZT$ is maximized. . . . .	134
Figure 5-6 : Depiction of the expected measurement from the scattering-parameter experiment depicted in Figure 5-5. . . . .	134
Figure A-1 : Graphical representation of the phase-space density flow described by the Boltzmann Transport Equation (BTE). . . . .	140

# List of Tables

4.1	Translation of various heat-density generation rates calculated internally in the Sentarus Device thermodynamic model into the notation of this document. Note that the electron chemical potential $\mu_n$ is equivalent to $E_F - E_C$ . . . . .	119
-----	--	-----



# Chapter 1

## Introduction

Microscopic theories of electrical and thermal conduction in solid-state systems typically separate the dynamics of the mobile electronic charge carriers from that of the underlying lattice of ions which produce the potential in which those carriers move. As we are well aware in the design of electronic devices, while both subsystems contribute to the conduction of heat, only the carriers contribute to the conduction of charge. That is to say, while the interactions of the two subsystems must be accounted for (typically through scattering), in an accounting sense the phonon lattice excitations contribute only to the flow of heat but the electronic subsystem contributes to the flow of both heat and electricity. This latter fact, that the thermal and electrical responses of any solid-state system are invariably linked as they share a single carrier distribution, is the origin of all the thermoelectric effects discussed herein and therefore underlies all practical devices based on these effects.

The same separation of dynamics is also the starting point for the theoretical description of all solid-state devices, including the design and simulation of their thermal operating characteristics. Nevertheless, the formalisms typically employed in the design of such devices as semiconductor lasers, bipolar-junction transistors (BJTs), and field-effect transistors (FETs) differ from those typically used to explain thermoelectricity. Since the phenomena involved can be witnessed in the same semiconductor material systems, it should not be surprising that a unified formalism can be found which describes both thermoelectrics and other practical solid-state devices.

The most common picture employed to describe thermoelectricity is conceptually rooted in irreversible thermodynamics, and abstracts away the microscopic dynamics of semiclassical Boltzmann transport into transport coefficients. Meanwhile, the drift-diffusion equations remain popular for the description of transistors and lasers, even in devices whose design merits thermal management.

In this thesis, we describe a *generalized drift-diffusion* framework which permits a microscopic description of thermoelectric phenomena and of course retains the ability to explain other device phenomena. This framework could be used to describe and design thermoelectric elements, which because of the recent efforts to improve effective material parameters by introducing nano-scale features [1], increasingly resemble micro-electronic devices. Additionally, this framework could be useful in the design of transistors and lasers for which thermal management is key to improved performance, but which have required the use of a separate framework to examine the possible benefit of incorporating thermoelectric effects directly into device design [2].

We begin by presenting a brief overview of thermoelectric phenomena and the typical approach to their description, including the typical method of incorporating microscopic physics. We then outline some of the applications of thermoelectric materials and briefly mention the importance of thermal management in certain modern micro-electronic devices.

## 1.1 Thermoelectric Phenomena

Our first goal is to describe the thermoelectric effects phenomenologically in order to quantify and communicate their practical value, and to generally motivate their microscopic description.

Applying a small temperature difference across a sample gives rise to a small longitudinal voltage under electrical open-circuit conditions by what is known as the Seebeck effect (see Figure 1-1). The ratio of the small resultant voltage to the initial

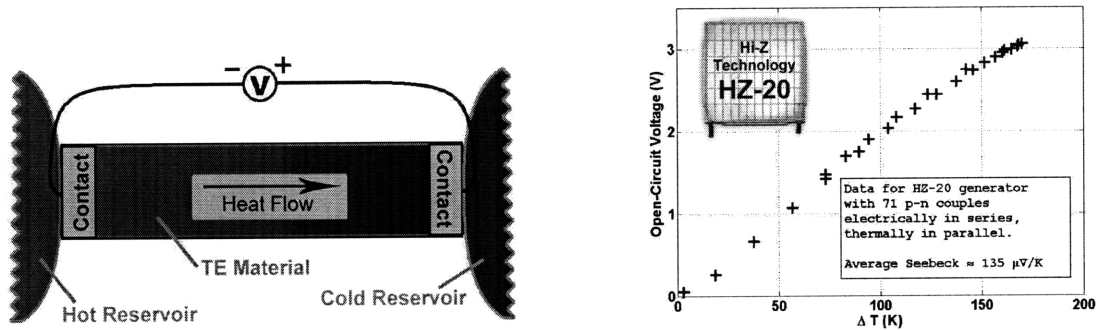


Figure 1-1: Left: Phenomenological depiction of the Seebeck effect. Right: Plot of open-circuit voltage with temperature difference across a commercial thermoelectric generator consisting of 142 elements connected electrically in series and thermally in parallel. This measurement was made in a demonstration power-generator built by the author. Note that the observed “average Seebeck” above is diminished from the material property value because of thermal parasitics.

small temperature difference is known as the Seebeck coefficient:

$$\frac{\Delta V}{\Delta T} \equiv \alpha \quad (1.1)$$

The Seebeck coefficient carries units of [Volts/Kelvin]=[Joules/Kelvin]/Coulomb], and though justification is left to Section 1.2.2, it may be thought of generally as the entropy transported per coulomb of charged carriers contributing to conduction. Note that we have chosen to define  $\alpha$  from differential quantities because it is a linear transport coefficient. The same is true of all transport coefficients referred to in this chapter, even if the smallness of the quantities whose ratios they reflect is omitted.

Another thermoelectric effect, the Peltier effect, occurs when an electrical current is passed through an interface between different materials in an isothermal system, causing the junction to experience cooling or heating depending on the sign of the current (see Figure 1-2). The ratio of this heat production (taking care to include only the contribution  $\propto I^1$  and to omit the Joule heating from any contact resistance  $\propto I^2$ ) to the current required to produce it defines the Peltier coefficient:

$$\frac{Q_{gen}}{I} \equiv \Pi_2 - \Pi_1 \quad (1.2)$$

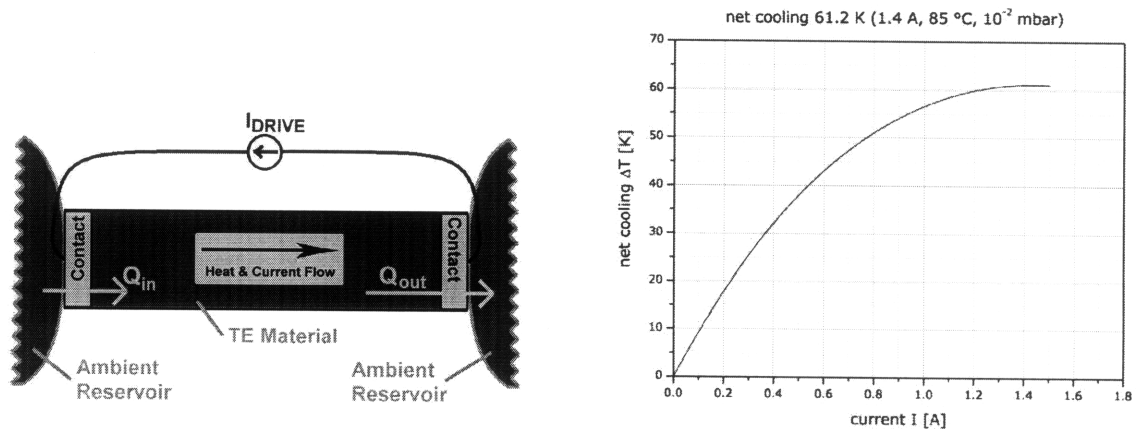


Figure 1-2: Left: Phenomenological depiction of the Peltier effect. Right: Plot of steady-state  $\Delta T$  across a commercial thermoelectric cooler versus current  $I$  required to maintain it. This Micro-Pelt MPC-D303 has 8 elements electrically in series and thermally in parallel. Plot taken from ([http://www.micropelt.com/products/peltier\\_cooler.php](http://www.micropelt.com/products/peltier_cooler.php)). Note that while the Peltier effect is defined under isothermal conditions, driving a current through a TE material as in the figure at left will result in heat transport from the left contact to the right contact (for positive  $\Pi$ ), giving rise to a temperature difference  $\Delta T$ . In steady-state, this rate of heat transport exactly cancels the rate of ordinary heat conduction giving rise to a maximum  $\Delta T$  as shown in the plot at right and derived in Section 1.2.1.



The Peltier coefficient carries units of [Watts/Amp]=[Volts]=[Joules/Coulomb], and may be considered the energy transported per coulomb of contributing carriers.

The Peltier and Seebeck effects were first related by Lord Kelvin in 1858 via arguments of thermodynamic reversibility which he himself noted were flawed and could be made to generate spurious results [3]. Nevertheless the relationship between the Seebeck and Peltier coefficients, which held water empirically and were theoretically justified from time-reversal symmetry nearly a century later by Lars Onsager (see Section B.2), bears the name of the second Kelvin relation:

$$\Pi = \alpha T \tag{1.3}$$

## 1.2 Thermoelectric Devices

### 1.2.1 Bulk Material Analysis

For thermoelectric materials, it is possible to define a useful dimensionless figure-of-merit,  $ZT$ :

$$ZT \equiv \frac{\alpha^2 \sigma}{\kappa} T = \frac{\alpha^2 \sigma}{\kappa_{ph} + \kappa_e} T \tag{1.4}$$

where  $\sigma$  and  $\kappa$  are the total electrical and thermal conductivities respectively and the subscripts refer to the phonon and electron contributions to the material's thermal conductivity respectively. For the present bulk material model, we presume the material parameters to take on a single value over the extent of the thermo-element, an assumption which we will break when we introduce the full thermodynamic Onsager model below (see Section 1.2.2).

The figure-of-merit defined above is chosen so that the primary performance parameters of practical devices are simple, monotonically increasing functions of  $ZT$ , with all dependences on the Seebeck ( $\alpha$ ) or conductivities ( $\sigma$ ,  $\kappa$ ) found in the combination  $Z$ . For this reason, increasing the  $ZT$  at relevant operating temperatures has become the central objective of the materials-development segment of the thermoelectric research community. Our discussion in support of this point follows closely

the text of Nolas, Sharp, and Goldsmid [4].

## Thermoelectric Generators

The energy-conversion efficiency of a thermoelectric generator (TEG), as with any heat engine, is given by the ratio of the rate it performs work at to the thermal input power required to do so. For a TEG this translates to the ratio of the electrical power it is able to deposit in a resistive load to the thermal power flowing into the TEG at the hot reservoir. For concreteness, we may imagine a modified version of the general setup from Figure 1-1 which replaces the voltmeter with a load resistor  $R_{load}$  equal to  $\beta$  times the internal resistance of the thermoelectric element, and specifies  $L$  and  $A$  to be the length and cross-sectional area of the device respectively. Noting that the thermally conducted power can be simply separated into a part determined by the thermal boundary conditions and another part due to the diffusion of Joule heat in the element, we have:

$$\begin{aligned}
 \eta &= \frac{P_{electrical,load}}{P_{thermal,hot-side}} \\
 &= \frac{I^2 \cdot R_{load}}{\kappa \nabla T A + \alpha T_H I} \\
 &= \frac{I^2 \cdot \beta R_I}{\kappa \frac{\Delta T}{L} A - \frac{1}{2} I^2 R_I + \alpha T_H I} \\
 &= \frac{\alpha^2 (\Delta T)^2}{R_I \cdot \left[ \kappa \frac{\Delta T}{L} A - \frac{1}{2} \left( \frac{\alpha (\Delta T)}{(1+\beta) R_I} \right)^2 R_I + \alpha T_H \left( \frac{\alpha (\Delta T)}{(1+\beta) R_I} \right) \right]} \cdot \left( \frac{\beta}{(1+\beta)^2} \right)
 \end{aligned} \tag{1.5}$$

Now if we choose to operate the generator at the load required for maximum power, we find that the maximum of  $\frac{\beta}{(1+\beta)^2}$  gives us  $\beta \rightarrow 1$  or simply  $R_{load} \rightarrow R_I$ . In this

case, the efficiency reduces to:

$$\begin{aligned}
\eta_P &= \frac{\alpha^2(\Delta T)^2}{R_I \cdot \left[ \kappa \frac{\Delta T}{L} A - \frac{1}{2} \left( \frac{\alpha(\Delta T)}{2R_I} \right)^2 R_I + \alpha T_H \left( \frac{\alpha(\Delta T)}{2R_I} \right) \right]} \cdot \frac{1}{4} \\
&= \frac{\alpha^2(\Delta T)^2}{\frac{\kappa}{\sigma}(\Delta T) - \frac{1}{8}\alpha^2(\Delta T)^2 + \frac{1}{2}\alpha^2 T_H(\Delta T)} \cdot \frac{1}{4} \\
&= \frac{T_H - T_C}{\frac{1}{Z} - \frac{1}{8}(\Delta T) + \frac{1}{2}T_H} \cdot \frac{1}{4} \\
&= \frac{1}{4} \cdot \eta_C \cdot \left( \frac{1}{ZT_H} - \frac{\eta_C}{8} + \frac{1}{2} \right)^{-1}
\end{aligned} \tag{1.6}$$

where  $\eta_C$  is the Carnot efficiency. Note that in the case where thermal conduction at the hot side approaches the electrical open-circuit value under the same temperature boundary conditions, the efficiency at max power is a simple linear function of  $ZT_H$ :

$$\eta_P \longrightarrow \frac{\eta_C}{4} \cdot ZT_H \tag{1.7}$$

### Thermoelectric Coolers

A similar analysis may be performed for refrigerating devices, built on the Peltier effect instead of the Seebeck effect. For such devices, one may derive expressions for the coefficient of performance ( $\phi$ ) between reservoirs of specified temperature and the maximum achievable  $\Delta T$ .

For a single uncouple, we may express  $\phi$  as the ratio of the heat flowing out of the cold reservoir through the thermo-element to the electrical power required to do so. When we describe the device as pumping heat from a cold reservoir at  $T_C$  to a hot reservoir at  $T_H = T_C + \Delta T$ , the net heat being pumped out is given by the peltier heat flow out minus the incoming heat from conduction:

$$Q_C = (\alpha T_C)I - \kappa \frac{\Delta T}{L} \cdot A - \frac{1}{2}I^2 R \tag{1.8}$$

where  $A$  and  $L$  are the cross-sectional area and length of the element respectively and  $R$  is the electrical resistance. Note that the conducted heat has nicely separated

into a component independent of current  $I$  and half of the total Joule-heating contribution. We can also simply describe the input electrical power as the sum of the Joule heating and any difference in carrier electrochemical potential that arises from the Seebeck effect across the reservoirs.

$$W = I(IR + \alpha(\Delta T)) \quad (1.9)$$

We pause here to interpret the above equation; we will return shortly to the question of  $\phi(T_C, T_H)$ . To run a current in the direction needed to carry heat away from the cold reservoir, one must perform work to push carriers upstream against a density gradient, which even for neutral particles would be entropically disfavored, but moreover is energetically disfavored because of the Coulomb repulsion of more carriers on the destination end of the thermo-element. This interpretation of the Seebeck effect will prove useful again in Chapter 3.

The ratio gives us  $\phi$ :

$$\phi = \frac{(\alpha T_C)I - K(\Delta T) - \frac{1}{2}I^2 R}{I^2 R + \alpha(\Delta T)I} \quad (1.10)$$

where  $K$  is the thermal conductance. Note that if we examine the limit of removing all sources of irreversibility by taking  $K \rightarrow 0$ ,  $R \rightarrow 0$ , we arrive at the Carnot limit  $\phi \rightarrow T_C/\Delta T$  [4]. As in the case of the generator, the maximum efficiency condition corresponds to a different operating point in general than the maximum power condition. Since the thermal conductance of the sample is taken as constant, to find the maximum achievable  $\Delta T$ , we need only choose the current to maximize the Peltier heat-pumping power minus Joule heat diffusion:

$$I_Q \equiv I|_{Q_C=Q_{C,max}} = \frac{\alpha T_C}{R} \quad (1.11)$$

then solve for the  $\Delta T$  which permits a maximum of  $Q_C = 0$ :

$$(\Delta T)_{max} = \frac{\frac{1}{2}I^2 R - \alpha T_C I}{-K} = \frac{1}{2} Z T_C^2 \quad (1.12)$$

## 1.2.2 Thermodynamic Onsager Analysis

As seen above, for devices of practical interest, the device performance parameters are monotonically increasing functions of  $ZT$ , such that from a materials-development perspective, it often suffices to follow strategies which pursue a higher figure-of-merit. Nevertheless, since material parameters ( $\alpha$ ,  $\sigma$ ,  $\kappa$ ) change with temperature, it is often useful to design a single thermo-element to operate effectively in its particular temperature regime. However, because of the need to match both electrical and thermal boundary conditions simultaneously [5, 6], since both carrier and energy density are conserved locally, this task of optimizing such inhomogeneous thermoelectric elements is not as simple as changing the material to maximize  $ZT$  at each location. Whether the approach taken is to change the dopant concentration or alloy composition continuously over the sample (known as functional grading), or to combine a small number of homogenous elements in parallel or series as needed (known as segmentation or cascading), the transport in such devices is poorly described by the simple model above.

Moreover, while a new generation of nano-structured thin-film thermoelectric materials has been developed to reach higher room-temperature  $ZT$ 's, the fabrication techniques required for these materials often exhibit drift, so that even elements designed to be homogeneous are often poorly described by the model expressed above.

Whether purposeful or accidental, modern thermoelectric elements are rarely well-described by bulk-element analysis. Instead, a single framework for the flow of charge and heat in thermoelectrics is needed to appropriately model and design such thermoelectrics. The thermoelectric Onsager Relations, an expression of the irreversible thermodynamics of charge and heat flow, constitute such a model and are derived here. The discussion presented here follows closely the Callen's text on thermodynamics [3].

We begin by expressing the rate of local entropy-density generation (not to be confused with accumulation) in terms of gradients of the temperature  $T$  and electro-

chemical potential  $\mu$  and the fluxes of energy  $\vec{J}_U$  and carriers  $\vec{J}_N$ :

$$\dot{s} = \left( \nabla \frac{1}{T} \right) \cdot \left( \vec{J}_U \right) + \left( \nabla \frac{\mu}{T} \right) \cdot \left( -\vec{J}_N \right) \quad (1.13)$$

where  $\mu$  is the electro-chemical potential. To motivate this equation, we recall from the micro-canonical ensemble of equilibrium statistical mechanics that the inverse-temperature  $1/T$  is defined as the partial derivative of the entropy with respect to the overall system energy:  $1/T = \frac{\partial S}{\partial U}$ . Similarly, the electro-chemical potential is given by  $\mu/T = \frac{\partial S}{\partial N}$ . If we consider a nonzero flux of energy  $\vec{J}_U$  to be indicative of bits of energy  $dU$  flowing from one small thermodynamic subsystem to an adjacent small subsystem with greater inverse-temperature, the overall entropy of the two combined subsystems will be increased by an amount proportional to the amount of energy transferred ( $dU$ ) and the inverse-temperature difference ( $\Delta(1/T)$ ). In this process, entropy has been generated. Using a similar logic for the flow of particles, we may see the above equation as intuitively expressing how the rate of local entropy generation can be expressed in terms of the flow of various extensive quantities. For a more formal derivation of the above equation and other results related to the Onsager Relations, see Appendix B.

Returning to (1.13), if instead of the flux of energy, if we choose to consider only the flux of heat, we have:

$$\vec{J}_Q = T \vec{J}_S = \vec{J}_U - \mu \vec{J}_N \quad (1.14)$$

and the expression becomes:

$$\dot{s} = \left( \nabla \frac{1}{T} \right) \cdot \left( \vec{J}_Q \right) + \left( \frac{1}{T} \nabla \mu \right) \cdot \left( -\vec{J}_N \right) \quad (1.15)$$

Note that  $\mu$  is the electro-chemical potential, which is the sum of the chemical potential for the charge carriers (electrons here) and their electrostatic potential energy.

Making use of the Onsager Reciprocity, we may then write the fluxes in terms of

the affinities, assuming a linear, memoryless system:

$$\begin{aligned} -\vec{J}_N &= L_{11} \frac{1}{T} \nabla \mu + L_{12} \nabla \frac{1}{T} \\ \vec{J}_Q &= L_{21} \frac{1}{T} \nabla \mu + L_{22} \nabla \frac{1}{T} \end{aligned}$$

where  $L_{12} = L_{21}$  for  $\vec{B} = 0$ .

From here, one can set up various “experiments” by enforcing boundary conditions and defining the usual transport coefficients in terms of the  $L$ ’s.

We define the electrical conductivity to be the ratio of the electrical current  $q\vec{J}_N$  to the gradient in the electro-chemical potential per unit charge  $\nabla\mu/q$  for  $\nabla T \rightarrow 0$ , where  $q$  is the charge of an electron. We define the thermal conductivity to be the ratio of heat flow  $\vec{J}_Q$  to the negative temperature gradient  $-\nabla T$  under conditions of no current flow  $\vec{J}_N \rightarrow 0$ . Likewise, we may set up the Seebeck and Peltier effects to find that:

$$\begin{aligned} -\vec{J}_N &= \frac{\sigma}{q^2} \nabla \mu + \frac{-T^2 \sigma \alpha}{q} \nabla \frac{1}{T} \\ \vec{J}_Q &= \frac{-\sigma \Pi}{q} \nabla \mu + (T^3 \sigma \alpha^2 + T^2 \kappa) \nabla \frac{1}{T} \end{aligned} \tag{1.16}$$

where the Onsager Reciprocity implies the Second Kelvin Relation:

$$\Pi = \alpha T \tag{1.17}$$

### 1.2.3 Microscopic Derivation of Transport Coefficients

Both of the models above describe thermoelectric devices and their underlying phenomena in terms of macroscopic transport coefficients without reference to their microscopic origins. Therefore, for these models to explain the performance of practical devices in terms of microscopic physics, we need only calculate the values of the transport coefficients (at each point in space for the thermodynamic Onsager model, but only once for the bulk picture) and insert them into the analyses above.

The microscopic calculation of electrical, and in particular, thermal conductivities

in recently developed materials with nanoscale features remains a topic of active research [7]. Nevertheless, these values in many materials of importance can at least be approximated with very simple models. To get more accurate predictive estimates of the conductivities, one may choose to view the simple model as an approximation of a more complex model, which when examined further results in a more accurate answer.

For example, the lattice contribution to thermal conductivity can be given by

$$\kappa_{ph} = \frac{1}{3} C_V c l \quad (1.18)$$

where  $C_V$  is the heat capacity at constant volume for the lattice (which can be estimated in terms of microscopic parameters by the Einstein-Debye model),  $c$  is the sound-speed, and  $l$  is the average mean-free-path [4]. Instead of assuming that all of the phonons are well-described by their average mean-free-path, if we consider the contributions to the heat capacity from phonons of various mean-free-path lengths, we can achieve a closer approximation. This is particularly useful when examining materials with defects specifically introduced to reduce the mean-free-path of a particular group of phonons [7].

Likewise, the electrical conductivity of a semiconductor is easily described in terms of the Drude model:

$$\sigma = qn\mu = qn \left( \frac{q\tau}{m^*} \right) \quad (1.19)$$

but can be further examined by considering contributions from various energies (a procedure undertaken in Section 2.4.1), or even wavevector.

Once the electrical conductivity has been calculated, the Law of Wiedemann-Franz [8] can be used to estimate the electronic contribution to thermal conductivity as:

$$\kappa_e = \frac{3}{2} \left( \frac{k_B}{q} \right)^2 T \sigma \quad (1.20)$$

to complete our microscopic estimation of the conductivities.

On the other hand the third transport coefficient which is present in the bulk and



Onsager models, the Seebeck coefficient, requires a longer calculation. In a sense, one of the main thrusts of this thesis is to develop a fuller model of this calculation in the same sense as the fuller models of the conductivity calculations above. The recipe in this section expresses contributions to the Seebeck coefficient from electronic carriers of various energies, while the model in Chapter 2 expresses the contributions from carriers of various wavevectors.

While the derivation presented here has been the basis for numerous theoretical investigations [9, 10, 11, 12, 2], our treatment here will primarily follow Section 1.3 of Pipe's doctoral thesis and the publications cited therein [2, 13], with the notable exception that we only consider unipolar semiconductors in this thesis. To solve for the Seebeck coefficient in terms of microscopic parameters, we begin with the Boltzmann Transport Equation (BTE) under the Relaxation-Time Approximation (RTA). In steady-state, taking the phase space density  $f$  to be a function of both space and wavevector, and taking the momentum-relaxation time  $\tau$  to be a function of wavevector, we have:

$$\frac{\hbar\vec{k}}{m^*} \cdot \nabla f + \frac{q\vec{E}}{\hbar} \cdot \nabla_k f = -\frac{f - f_0}{\tau} \quad (1.21)$$

If we make the assumption of spherical parabolic bands with dynamical effective mass  $m^*$ , we know that in equilibrium the electronic carriers occupy states with density given by the Fermi-Dirac distribution:

$$f_0(E) = \left[ e^{\frac{E + (E_C - E_F)}{k_B T}} + 1 \right]^{-1} \quad (1.22)$$

where

$$E = \frac{\hbar^2 k^2}{2m^*} \quad (1.23)$$

Here we take the energy  $E$  as zero at the band-edge and explicitly permit gradients in the band-edge and hence in the vacuum level, thereby circumventing a later inaccurate assumption of the electro-chemical potential being roughly equal to the chemical potential that was made in [2].

To first order in the real- and momentum-space gradients, we then have

$$\nabla f \approx \nabla f_0 = \frac{\partial f_0}{\partial E} \left( \frac{-(E + E_C - E_F)}{T} \nabla T + \nabla(E_C - E_F) \right) \quad (1.24)$$

and

$$\nabla_k f \approx \nabla_k f_0 = \frac{\partial f_0}{\partial E} \hbar \frac{\vec{k}}{m^*} \quad (1.25)$$

Substituting these expressions into the BTE and recalling that  $\vec{E} = -\nabla E_C/q$ , we can solve for the deviation of  $f$  from equilibrium:

$$\tau \frac{\partial f_0}{\partial E} \frac{\hbar \vec{k}}{m^*} \cdot \left[ \frac{-(E + E_C - E_F)}{T} \nabla T - \nabla E_F \right] = f_0 - f \quad (1.26)$$

Using this expression for  $f$ , we can express the current density in terms of a contribution proportional to  $\nabla T$  and another proportional to  $\nabla E_F$ , then identify the coefficients as  $\sigma\alpha$  and  $\sigma$  respectively.

$$\begin{aligned} \vec{J} \cdot \hat{z} &= 2q \int_{\mathbb{B}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k_z}{m^*} f(\vec{k}) \\ &= 2q \int_{\mathbb{B}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k_z}{m^*} (f - f_0) \\ &= -2q \int_{\mathbb{B}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k_z}{m^*} \tau(\vec{k}) \frac{\partial f_0}{\partial E} \frac{\hbar \vec{k}}{m^*} \cdot \left[ \frac{-(E + E_C - E_F)}{T} \nabla T - \nabla E_F \right] \\ &= -2q \int_{\mathbb{B}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar k_z}{m^*} \tau(\vec{k}) \frac{\partial f_0}{\partial E} \frac{\hbar \vec{k}}{m^*} \cdot \left[ \frac{-(E + E_C - E_F)}{T} \nabla T - \nabla E_F \right] \end{aligned} \quad (1.27)$$

Let us assume that  $\tau(\vec{k}) = \tau_0 [E(\vec{k})]^r$  and take the gradients to be in the same ( $\hat{z}$ ) direction as we are asking about the current in. Converting this expression into an integral over  $E$ , we can write the result in terms of defined integrals as follows:

$$\vec{J} \cdot \hat{z} = \frac{-2q\sqrt{2m^*}\tau_0}{3\pi^2\hbar^3} \left[ \left( \frac{\nabla T}{T} \right) ((E_F - E_C)\mathbb{I}_r - \mathbb{I}_{r+1}) + (-\nabla E_F) \mathbb{I}_r \right] \quad (1.28)$$

where

$$\mathbb{I}_\gamma \equiv \int_0^\infty dE \left[ E^{\frac{3}{2}+\gamma} \frac{\partial f_0}{\partial E} \right] \quad (1.29)$$

Which leads us to expressions for the conductivity and Seebeck coefficient in terms of microscopic parameters and an integral over energy.

$$\sigma = \frac{(\vec{J} \cdot \hat{z})|_{\nabla T \rightarrow 0}}{-\nabla E_F/q} = \left( \frac{-2q^2 \sqrt{2m^* \tau_0}}{3\pi^2 \hbar^3} \right) \mathbb{I}_r \quad (1.30)$$

and

$$\alpha = \frac{(\vec{J} \cdot \hat{z})|_{\nabla E_F \rightarrow 0}}{-\nabla T \cdot \sigma} = \frac{-1}{Tq} \frac{((E_F - E_C)\mathbb{I}_r - \mathbb{I}_{r+1})}{\mathbb{I}_r} \quad (1.31)$$

For the simplified case of the Boltzmann limit, where  $f_0(E) = \exp[-(E + E_C - E_F)/(k_B T)]$ , we have:

$$\begin{aligned} \frac{\partial f_0}{\partial E} &= \exp[-(E + E_C - E_F)/(k_B T)] \left( \frac{-1}{k_B T} \right) \\ \mathbb{I}_\gamma &= \frac{-e^{-(E_C - E_F)/(k_B T)}}{k_B T} \int_0^\infty dE E^{\frac{3}{2} + \gamma} \exp[-E/(k_B T)] \\ &= [-e^{-(E_C - E_F)/(k_B T)}] \Gamma\left(\frac{5}{2} + \gamma\right) (k_B T)^{\frac{3}{2} + \gamma} \end{aligned} \quad (1.32)$$

so that using the identity  $\Gamma(n + 1) = n\Gamma(n)$  and substituting in the expression  $n = N_C e^{-(E_C - E_F)/(k_B T)}$  for the carrier density in the Boltzmann limit, we arrive at:

$$\begin{aligned} \alpha &= \frac{-1}{Tq} (E_F - E_C) - (k_B T) \left( \frac{5}{2} + r \right) \\ &= \frac{k_B}{q} \left[ \frac{5}{2} + r - \log\left(\frac{n}{N_C}\right) \right] \end{aligned} \quad (1.33)$$

which is the common expression for the Seebeck coefficient in the Boltzmann limit.

Even outside the Boltzmann limit, it is possible to rewrite the expression for the conductivity from (1.30) in terms of an integral over the Fermi window:

$$\begin{aligned} \sigma &= \left( \frac{2q^2 \sqrt{2m^* \tau_0}}{3\pi^2 \hbar^3} \right) \int_0^\infty dE \left( -\frac{\partial f_0}{\partial E} \right) E^{\frac{3}{2} + r} \\ &\equiv \int_0^\infty dE \left( -\frac{\partial f_0}{\partial E} \right) \sigma'(E) \end{aligned} \quad (1.34)$$

where we may regard the preceding expression as the definition of the differential

conductivity  $\sigma'(E)$ . We may then also express the Seebeck coefficient in terms of  $\sigma'(E)$ :

$$\begin{aligned}\alpha &= \frac{-1}{Tq} \frac{\int_0^\infty dE \left(-\frac{\partial f_0}{\partial E}\right) ((E_F - E_C) - E) \sigma'(E)}{\int_0^\infty dE \left(-\frac{\partial f_0}{\partial E}\right) \sigma'(E)} \\ &= \frac{\langle E + E_C - E_F \rangle_{\sigma'(E)}}{Tq}\end{aligned}\tag{1.35}$$

where  $\langle \cdot \rangle_{\sigma'(E)}$  may be regarded as the differential-conductivity weighted average over the Fermi window. Because the Peltier coefficient is commonly considered to represent the average transport energy, here  $(E + E_C - E_F)$ , the above expression suggests that the quantity  $\sigma'(E)$  which encodes the relative importance of various energies to electrical charge conduction, also encodes the relative importance of various energies to the Peltier and Seebeck coefficients. For this reason, we refer to this derivation of  $\alpha$  as the differential-conductivity method.

This microscopic analysis of the Seebeck effect begins with the BTE and so therefore also carries the assumptions of semiclassical transport Appendix A, as do all other derivations beginning there including that in Chapter 2. However the combination of the Onsager thermodynamic model and the transport coefficient calculations we just performed also implicitly make an assumption about the homogeneity of a thermoelectric element. They essentially divide space into small regions for which the transport coefficients are calculated and combine them using the Onsager equations. The size of these regions therefore introduces another length-scale into the problem, which must be large enough for the divided chunks of space to obey the laws of thermodynamics and avoid the complications of microscopic defects and impurities yet small enough to permit the consideration of each region as homogeneous. For so-called functionally-graded materials, whose material parameters are changed continuously along the length of a thermoelectric element, or for thin films which operate under very large temperature gradients where the transport coefficients could change substantially on short length-scales, this choice must be made carefully.

With the real-space model in Chapter 3, we avoid this difficulty entirely by considering all quantities related to boundary conditions  $(T, E_F)$  in the microscopic

analysis and consider quantities like the electrical conductivity and Seebeck coefficient to be properties of the set of self-consistent solutions of the microscopic physical laws. Moreover we witness phenomena that does not manifest in the point-wise microscopic model, such as density variations [14] at the ends of even homogeneously-doped elements and the associated local changes in the band-edge profile.

### 1.3 Thermoelectric Applications

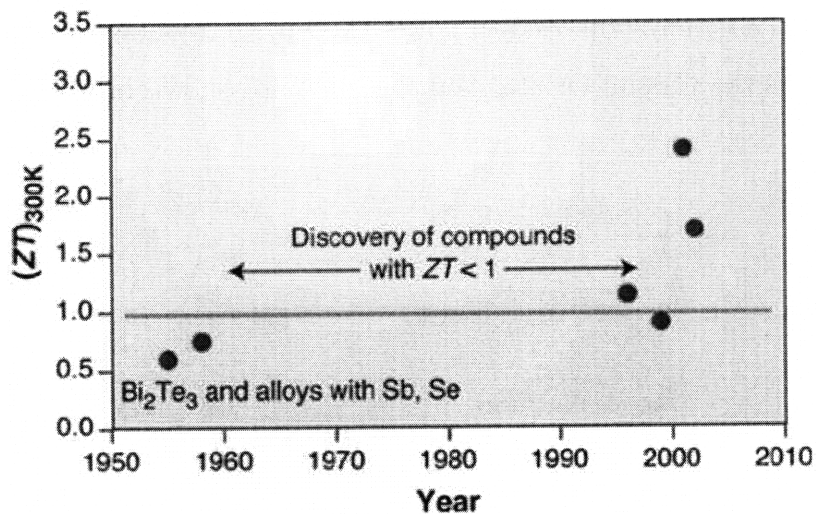


Figure 1-3: Room-temperature  $ZT$  achievements with time [1]. Text in the figure should read “Discovery of compounds with  $ZT > 1$ .”

For decades generators and refrigerators based on thermoelectric effects have been constrained to highly specialized applications which require the durability and compactness typically offered by these solid-state devices. The primary reason for this has been that the room-temperature figure-of-merit for available material systems remained stagnant at around  $ZT = 1$  (see Figure 1-3 [1]). As inhomogeneous and nano-structured thermoelectric elements are developed with higher room-temperature  $ZT$ 's, solid-state thermoelectric energy conversion may become viable options in application spaces historically filled by mechanical turbines and compressors [15]. Today's push for greater energy efficiency in automobiles may also be served by thermoelectric technologies [15]. As new technologies such as micro-scale wireless sensors

become more pervasive, the practical application of TEGs as compact ambient-energy harvesting elements [16] may also represent a new source of demand.

### 1.3.1 Direct Power-Generation

The efficiencies available with modern thermoelectric materials are insufficient for grid-scale power production. However, smaller niche applications have used TEGs for decades, including several of NASA's space probes [17]. Hundreds of compact, lightweight, reliable thermoelectric elements were used as part of a radioisotope thermal generator (RTG) system used to power the Voyager, Galileo, and Ulysses spacecraft, as well as the Cassini Orbiter launched in 1997, shown in Figure 1-4.

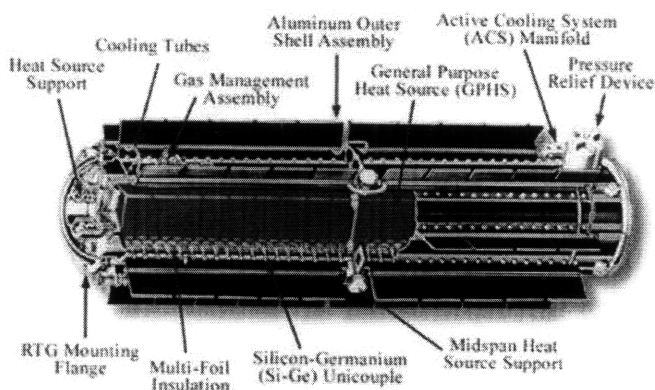


Figure 1-4: One of three RTGs used to power the Cassini Orbiter, launched 1997 [17].

Recent materials advances, both theoretical and experimental, have led to much optimism about the future use of TEGs in direct power generation. Nevertheless, thermoelectric technology remains drastically inefficient compared to modern gas-turbine combined-cycle power generators, which have recently passed the threshold of 60% thermal efficiency [18]. Even portable 10-kilowatt gasoline-powered generators can achieve thermal efficiencies of 25% [19]. With a typical hot-side temperature of  $T_H = 2000K$  and a cold-side temperature of  $T_C = 400K$ , for a TEG to achieve similar performance would require  $ZT \approx 400$  and  $ZT \approx 12$  respectively [20].

Even if materials with such spectacular  $ZT$ 's were made, practical power genera-

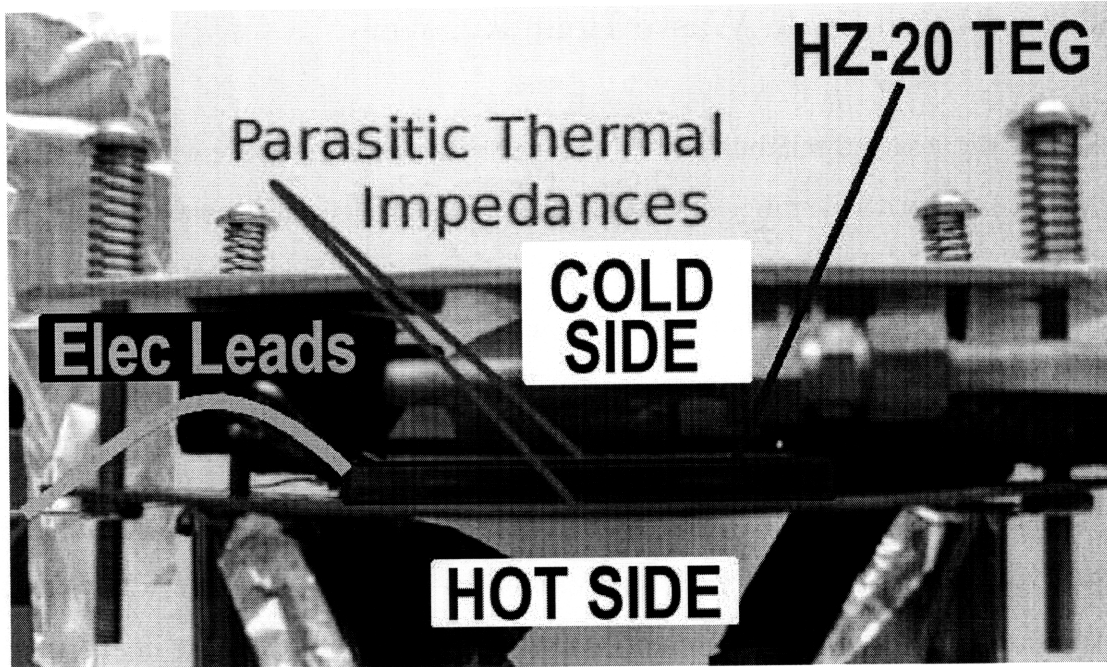


Figure 1-5: Demonstration thermoelectric generator with large thermal parasitics, built by the author.

tion still faces numerous challenges. In addition to the problem of ensuring that the parasitic thermal resistances in series with the generator do not absorb a substantial fraction of the temperature drop (see Figure 1-5), the issue of thermoelectric compatibility, analogous to the compatibility issue with multi-junction photovoltaics, must be addressed. Essentially, the generated current at the optimal power point must be matched for all segmented devices in series and the relative thermal impedances of various segments must permit the optimal temperature profile in order to achieve ideal segmented performance [5, 6]. For many large-scale applications therefore, not only must an arsenal of good thermoelectric materials which span the desired temperature regime be developed, but the individual materials must further satisfy the requirement of so-called thermoelectric compatibility [5, 6]. However if these issues are addressed and materials satisfying the necessary requirements are developed, an exceedingly large application space awaits.

### 1.3.2 Macro-Scale Waste Heat Recovery

The compactness and reliability of thermoelectric generators, along with their ability to operate across a wide range of heat qualities permits their application in waste-heat recovery systems. While virtually every modern machine could be thought of as a heat engine operating below Carnot efficiency, and therefore represents an opportunity to recover some of the wasted heat to improve overall efficiency, a few opportunities are sufficiently lucrative to be considered even with current thermoelectric technologies.

A 2004 proposal, funded by the US Department of Energy, was undertaken by a joint team from BMW, BSST, and Visteon in collaboration with academic and government laboratories, to attempt to achieve a 10% improvement in the overall efficiency of a modern passenger vehicle [21]. The strategy is to sap waste heat from the exhaust system (see Figure 1-6 and Figure 1-7) and have a TEG convert it into enough electrical power to satisfy the majority of the power load normally placed on the alternator. The additional thermal impedeance in the exhaust line should also reduce warm-up time and thereby improve vehicle fuel efficiency further in cold weather.

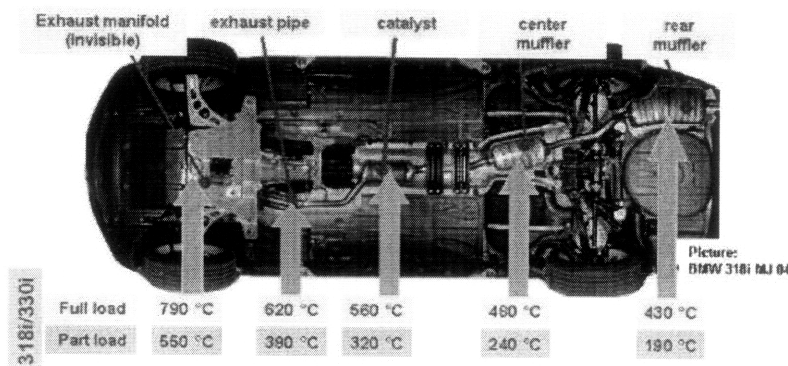


Figure 1-6: Operating temperature profile for exhaust system of a BMW passenger vehicle [21].



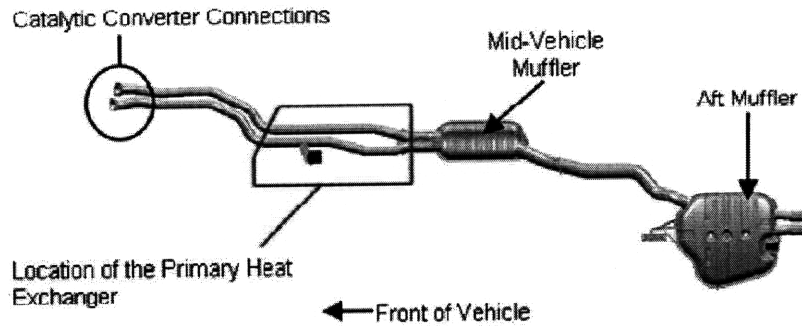


Figure 1-7: Proposed site for waste-heat recovery in an automotive TEG system. Image from (<http://www.bsst.com/automotive-waste-heat-recovery.php>).

### 1.3.3 Micro-scale Wireless Sensor Power

As numerous applications for small electronic sensors develop, a need for energy scavenging devices has emerged. Thermoelectric technology is particularly well-suited to this environment for numerous reasons [16], including the durability and size of the devices as well as the ubiquity of their power source. Competing technologies often rely on unreliable features of the ambient environment, such as stray radio-frequency power from electromagnetic communications or solar power. Background rf power in urban areas is typically around a few  $\mu\text{W}/\text{cm}^2$ , below many current sensors' requirements. Meanwhile  $\text{mm}^3$ -scale sensors could easily find shadows if released outdoors, and would not have sufficient incoming power density if released indoors. Vibrational-energy scavenging devices are hurt by the rising fundamental frequency of small volumes, and hence become less effective for small sensors. Thermoelectrics, on the other hand, scale well with size and could make use of very common temperature differences between surfaces on which such  $\text{mm}^3$ -scale sensors might land and the ambient air. While issues with the stepping-up of DC voltages to usable levels for electronic sensors remain, since there is no fundamental limit to the efficiency of DC-DC electrical power conversion, thermoelectricity still looks well-poised to compete in the micro-scale wireless sensor power application space.

### 1.3.4 Cooling and Temperature-Stabilization

The Peltier effect can be exploited to pump heat against the direction of natural diffusion using electrical power. Since such devices use the dense gas of mobile electrons in the thermoelectric material as the working fluid, they offer environmental benefits over traditional compressor-based heat pumps found in commercial and industrial refrigerators [15]. Also, since such devices can be miniaturized without the need to design and construct assemblies of miniaturized moving parts, their employment becomes more attractive at smaller length-scales. Indeed, there is strong interest [22] in expanding the use of external thermoelectric devices to stabilize micro-electronic components [15], which have recently faced an adverse Moore's Law of increasing power dissipation density due to increased functionality per unit volume.

## 1.4 Thermal Self-Management Applications

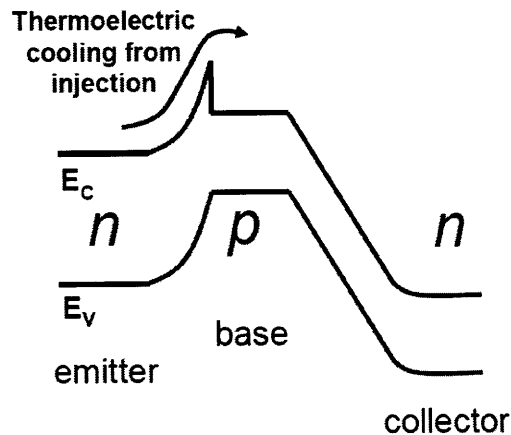


Figure 1-8: Consideration of internal Peltier cooling at the base-emitter junction of heterojunction bipolar transistors (HBTs) could lead to improved device designs [2].

In numerous applications, the heat produced by active micro-electronic devices degrades their operation and requires cooling. In some devices, the heating problem is extremely localized. For example, high temperatures at the base-emitter junction of BJTs and HBTs leads to noise in amplifiers made from them and high temperatures in

the active region of laser diodes leads to the strengthening of numerous non-radiative carrier recombination pathways that increase threshold current and reduce efficiency. While in many such cases it will remain attractive to cool such devices externally, when the heating problem is localized on the nano-scale, such as in the examples above, the consideration of thermoelectric effects in the original device design (see Figure 1-8) could enable the design of self-cooling devices, such as the ICICLE laser [2] shown in Figure 1-9.

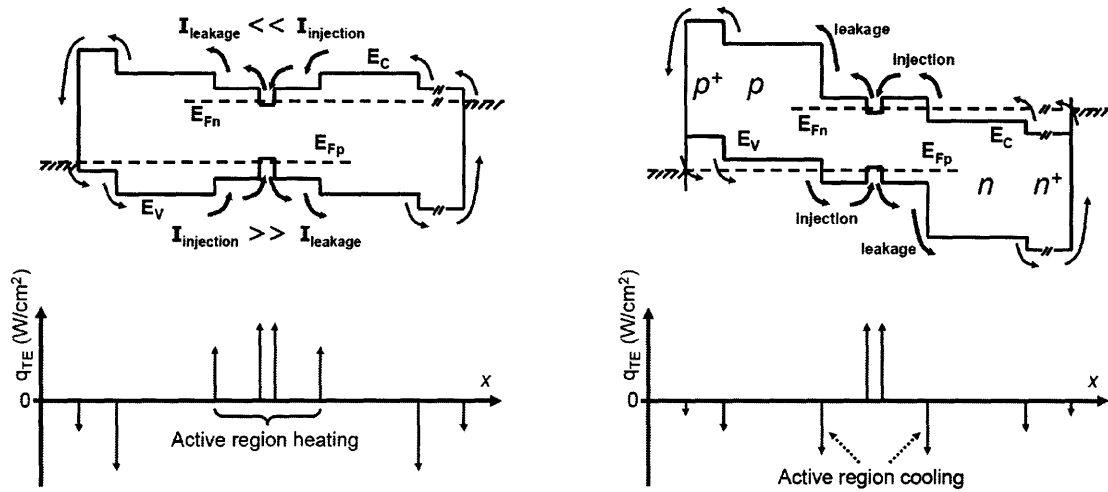


Figure 1-9: By examining the Peltier heating and cooling present in typical diode lasers, whose band diagram and Peltier heat generation profile shown left, Pipe proposed a design [2] for the ICICLE (Injection Current Internally Cooled Light Emitter), depicted at right.

## 1.5 Thesis Outline

In Chapter 1 we have presented the most common formalism for describing thermoelectric effects, and have motivated the development of a unified formalism for thermoelectric effects and micro-electronic device design. In the subsequent chapters, we present the generalized drift-diffusion picture which accomplishes this task, and begin to apply it to some specific examples.

In Chapter 2 we present the picture of semi-classical thermoelectric transport in reciprocal (or momentum) space. The differential conductivity approach related

in Section 1.2.3 describes thermoelectric transport primarily in energy-space, finally writing the Seebeck coefficient as a ratio of integrals over  $E$ . By examining the distribution function before performing the non-invertible map  $\vec{k} \rightarrow E(\vec{k})$  on the independent variable, we can gain intuition about the microscopic origin of thermoelectric effects not previously available to us and, as with the conductivity calculations in Section 1.2.3, extend the model to more accurately describe real materials.

In Chapter 3 we present the full description of the generalized drift-diffusion equations for thermoelectric transport. We begin with the simplifying assumption of a single energy-independent momentum-relaxation time, and use the framework to explain the Seebeck effect for an extrinsic semiconductor with spherical parabolic band-structure in the Boltzmann limit. Although the quantitative results match that of the differential-conductivity model, and agreement extends to the case of relaxation time with power-law dependence on momentum, this real-space model nonetheless offers a qualitatively new picture with particular implications for metal-semiconductor junctions at the contacts of a TEG as well as the physical interpretation of the analytical result.

In Chapter 4 we begin to use the framework to simulate Peltier heat exchange in micro-electronic devices and analyze inhomogeneous thermoelectric elements. As a first step in this process, we transition from using the relatively simple computational tools (developed by the author) which we used to describe the physics in preceding chapters to a commercial device simulator which incorporates a broader set of physical effects. We validate the new simulator by replicating the Seebeck and Peltier effects from Section 3.2 and Section 3.6, then move on to examine the  $p$ - $n$  junction.

In Chapter 5 we briefly outline the future research directions which the framework presented in this document suggests. We include the specific experimental prediction of a new mechanism for directly measuring the scattering exponent in a thermoelectric material by Atomic Force Microscopy as well as initial attempts to reconcile our numerical models with an unpublished observation by Lürßen and Ram of internal Peltier cooling in an  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$  tunnel diode by thermo-reflectance microscopy.

# Chapter 2

## Reciprocal-Space Picture of Thermoelectric Transport

### 2.1 Introduction and Overview

In this chapter, we seek a fuller description of the microscopic transport processes relevant to the thermoelectric effects than the differential-conductivity approach provided in Section 1.2.3.

In the differential-conductivity calculation of  $\alpha$  and  $\sigma$ , although our initial expressions were written in terms of integrals over reciprocal-space, we eventually reduced them each to a single integral over the energy  $E$ . Here we seek to examine a closely-related set of microscopic transport coefficients without performing the simplifying non-invertible map  $\vec{k} \rightarrow E(\vec{k})$  on the independent variable space. In doing so, we find intuitive descriptions of the transport coefficients which will help us reconstruct the thermoelectric effects microscopically in Chapter 3.

In Section 2.2 we motivate the new choice of variables which distinguish our transport calculations from the differential-conductivity approach. In Section 2.3 we begin our analytical derivation of transport coefficients by solving the Boltzmann Transport Equation under the Relaxation Time Approximation and graphically interpreting the expressions for the distribution function under electric potential, density, and temperature gradients. In section Section 2.4 we then use these results to analytically

express the conductivity, diffusivity, and Soret coefficient for cases of both energy-independent and energy-dependent relaxation times. In Section 2.5 we extend this analysis beyond the Boltzmann limit to include doping densities typical of practical thermoelectric elements.

## 2.2 The New Choice of Variables

The generalized drift-diffusion model for microscopic thermoelectricity presented in this thesis differs from the combined differential conductivity and thermodynamic Onsager model (see Section 1.2.3) in its choice of both independent and dependent variables.

As we saw with the conductivity calculations in Section 1.2.3, the consideration of the dynamics of a transport process in a more-complex, higher-dimensional independent variable space can offer additional intuition for the associated transport coefficients. In the case of lattice thermal conductivity, we chose to examine the contributions to the heat capacity from collections of phonon modes with particular mean-free paths. In the case of electrical conductivity, we chose to examine the contributions to electrical current from various energy levels for a system under bias. In fact, as we will see in Section 2.4.1, this segregation of contributions to conduction from various energies can offer us a correction term in the case of the energy-dependent relaxation time ( $\tau_0 \rightarrow \tau(E)$ ). In the case of thermoelectric transport, we modify the prevailing [9, 10, 11, 12, 2] model for transport by replacing the contributions from electronic carriers of various energies  $E$  with the more complex consideration of which wave-vector  $\vec{k}$  gave rise to that value of energy  $E(\vec{k})$ .

While the choice to view transport in reciprocal space over energy space is primarily taken to develop an intuitive picture, we may eventually find this description useful for several reasons. In the semi-classical transport framework (see Appendix A), the wave-vector  $\vec{k}$  gives the complete quantum wavefunction of the independent charge-carrier at a given point in space. Therefore in this description all transition rates can in principle be calculated, and we need not keep such quantities as  $\tau(\vec{k})$  as conceptual

black boxes. Also, for useful devices, we are often interested in thermoelectric effects in places where the phase-space density of carriers is dramatically out of equilibrium. While our particular calculations of transport coefficients will not be valid in such a regime, the integrals we used to solve for them require only the semiclassical approximations to remain valid, and so we could readily examine such situations numerically given some approximation of the phase-space density.

We also choose to change the dependent variables which describe the system at any point in space from the temperature and electro-chemical potential  $(T, E_F)$  to the temperature, density, and electric potential  $(T, n, \phi)$  of the electronic carriers. This choice is made primarily because  $E_F$  at a given location is itself explicitly dependent on the local temperature, clouding our ability to describe the transport of carriers in terms of independent sources that “drive” them. The new choice is useful not only because it does permit description in terms of such independent sources, but moreover the new quantities chosen as independent variables are conceptually simple and permit the joint solution of these equations with other physical formalisms. In particular, the choice to use the density  $n$  and potential  $\phi$  instead of the electro-chemical potential  $E_F$  permits the system to be solved self-consistently with Maxwell’s equations without regard for the thermodynamic interpretation of these quantities. The true value of this choice will be revealed in Chapter 3, where we will find that  $T$ ,  $n$ , and  $\phi$  emerge naturally from the Momentum Balance Equation, and the inclusion of Maxwell’s equations offers us an intuitive picture of the Seebeck effect in real-space.

For now, however, it suffices to consider this choice of dependent variables as convenient and intuitive, particularly for the device physicist familiar with the popular drift-diffusion transport model.

## 2.3 The Relaxation-Time Approximation

In the semi-classical picture, the central equation describing the dynamics of a non-equilibrium system of mobile carriers is the Boltzmann Transport Equation (BTE). For the steady-state solutions we seek, a greatly-simplifying assumption known as the

Relaxation-Time Approximation (RTA) permits analytical solution. The difficulties of reconciling reversible underlying mechanical laws with irreversible macroscopic behavior (briefly mentioned in Appendix A) are conceptually swept under the rug by the RTA, the use of which constitutes the replacement of a system trajectory given as an average over motions following reversible mechanical laws by a statistical closure that gives rise to irreversible dynamics. In systems with large numbers of particles that are close to equilibrium, however, the approximation is a good one and permits us to write the steady-state BTE as:

$$\left[ \frac{\hbar}{m^*} \vec{k} \cdot \nabla + \frac{q}{\hbar} \vec{E} \cdot \nabla_k \right] f = S[f] \approx \frac{-(f - f_0)}{\tau} \quad (2.1)$$

where  $f_0$  is the equilibrium phase-space density given by the Fermi-Dirac occupancy of the electronic states in the material's bandstructure. In the equation above we have purposefully suppressed the dependences on  $\vec{x}$  and  $\vec{k}$  because the relaxation time  $\tau$  is at times taken to be a function of various quantities, including  $E$  and  $\vec{k}$ . Recall that the energy is a function of wave-vector, and given the particles real-space location, the wave-vector completely describes a particle's quantum state, so that matrix elements for transitions can in principle be calculated and summed to find the characteristic relaxation time  $\tau$ . For this reason, we take as our general form  $\tau = \tau(\vec{k})$ . As usual, in steady-state, the density  $f = f(\vec{x}, \vec{k})$ .

Our goal, as mentioned briefly in our case for the use of  $(T, n, \phi)$  as dependent variables, is to calculate the deviation of the distribution of carriers from equilibrium in terms of these three independent driving gradients. Note that this information in effect includes, but is not limited to, the microscopic description of the transport coefficients relating particle flux to gradients in the dependent variables: the Soret coefficient  $S$ , the diffusion coefficient  $D$ , and the conductivity  $\sigma/q$ . These coefficients ask only for the relationship between the first moment of the distribution  $f$  and the associated source, whereas we seek to describe the deviation in every moment of the distribution.

It is worth mentioning that care should be taken in applying the results to particu-



lar physical configurations, such as is done in Chapter 3 and Chapter 4. Although we intend to apply these results to situations where more than one deviation is nonzero at a time, we will have independently considered the cases of nonzero electric fields  $\vec{E}$  and nonzero spatial gradients of  $n$  and  $T$ . Because we take the deviations from equilibrium to be small and we keep only the first-order contributions, the total deviation  $\delta f$  is simply the sum of the source-specific deviations ( $\delta f_{\vec{E}} + \delta f_{\nabla n} + \delta f_{\nabla T}$ ). To see why, consider that at each point in phase space, we are computing a Taylor expansion to first order in multiple orthogonal directions by independently summing their contributions at first order. Note again the value of our choice of  $(T, n, \phi)$ .

We now solve the BTE under RTA, interpret the results graphically, and give the basic expressions for the deviations due to the gradients in  $(T, n, \phi)$ . We leave the intuition about specific sources as well as calculations of transport coefficients in various cases of relevance to modern thermoelectricity to the subsequent sections.

Taking  $f \equiv f_0 + \delta f$ , for  $\vec{E} \neq 0$  and  $\nabla f_0 \neq 0$  respectively, unraveling the recursive definition from (2.1) gives:

$$\frac{q}{\hbar} \vec{E} \cdot \nabla_{\vec{k}} f = \frac{-\delta f}{\tau(\vec{k})} \implies f = f_0 + \frac{-q\tau(\vec{k})}{\hbar} \vec{E} \cdot \nabla_{\vec{k}} f_0 + \dots \approx f_0(\vec{x}, \vec{k} - \frac{q\tau(\vec{k})}{\hbar} \vec{E}) \quad (2.2)$$

$$\frac{\hbar}{m^*} \vec{k} \cdot \nabla f = \frac{-\delta f}{\tau(\vec{k})} \implies f = f_0 + \frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot \nabla f_0 + \dots \approx f_0(\vec{x} - \frac{\hbar\tau(\vec{k})}{m^*} \vec{k}, \vec{k}) \quad (2.3)$$

Now we turn our attention to interpreting these results.

As is commonly noted in texts on transport [23], the equation expressing  $f = f_0 + \delta f_{\vec{E}}$  constitutes the shifting of the distribution from mean  $\langle \vec{k} \rangle = 0$  to  $\langle \vec{k} \rangle = \frac{q\tau_0 \vec{E}}{\hbar}$  in the case of  $\tau(\vec{k}) \rightarrow \tau_0$ . The interpretation of this picture, closely approximated by the first order RTA result shown in Figure 2-1, is that the distribution of carriers in momentum-space looks identical to the case with zero  $\vec{E}$ , except that each particle has gained an average momentum equal to the amount of velocity change that the field would have imparted on it in time  $\tau$ . This is exactly the story described in the Drude model, so it is not surprising that the expression gives a density  $f$  whose

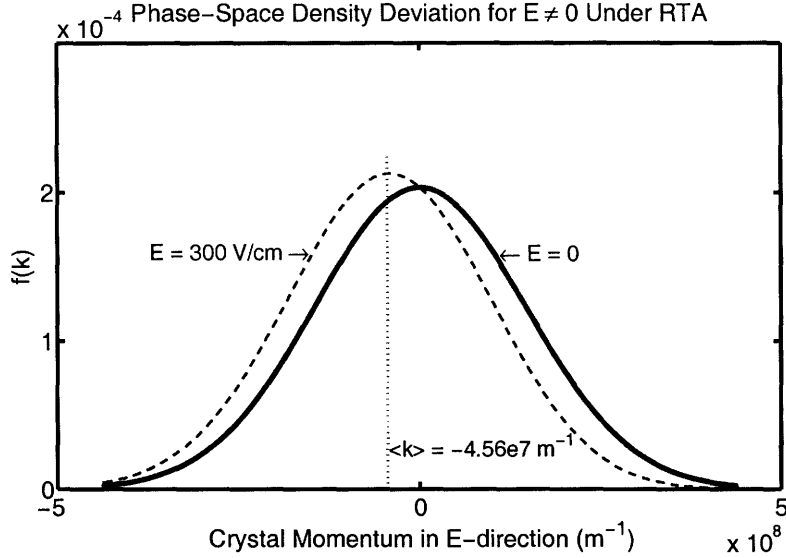


Figure 2-1: Plot of phase-space density as a function of  $\vec{E}$ -directed wave-vector for a distribution of negatively charged carriers under a positive electric field. The average wave-vector for the deviated distribution is  $\langle k \rangle = -4.56e7$ , indicating a flux of carriers to the left. Parameters for plot are as follows:  $\tau = 1\text{ps}$ ,  $m^* = 0.063m_e$ ,  $T = 300\text{K}$ , and  $n_D = 1e16 \text{ cm}^{-3}$ . The choice of ionized donor concentration puts the extrinsic semiconductor in the Boltzmann limit.

mean corresponds exactly to a current density  $\vec{J} = \sigma \vec{E}$  with  $\sigma = qn \left( \frac{q\tau}{m^*} \right)$  recovering exactly the Drude conductivity. The proper derivation of this result is presented in Section 2.4.1.

The interpretation of the expression for  $f$  in the presence of gradient quantities (plotted in Figure 2-2 and Figure 2-3) is complicated by the question of equilibrium. In a sense we have contradicted ourselves by assuming that  $f_0$  is the equilibrium density to which scattering processes eventually return the distribution, and yet permitting  $f_0$  to have  $\nabla T \neq 0$  or  $\nabla n \neq 0$ , which (with no other gradients nonzero) by definition makes the system out of equilibrium. What we really meant by  $f_0$  in the first place though, was the density to which scattering processes would on net return  $f$  in characteristic time  $\tau$ .

Let's examine more closely the case of the system under temperature gradient. Without loss of generality, we may take the direction of the gradient to be  $\hat{z}$ . For particles with zero momentum in the direction of the gradient, meaning those in the  $k_z = 0$  plane, the temperature which characterizes the Fermi function according

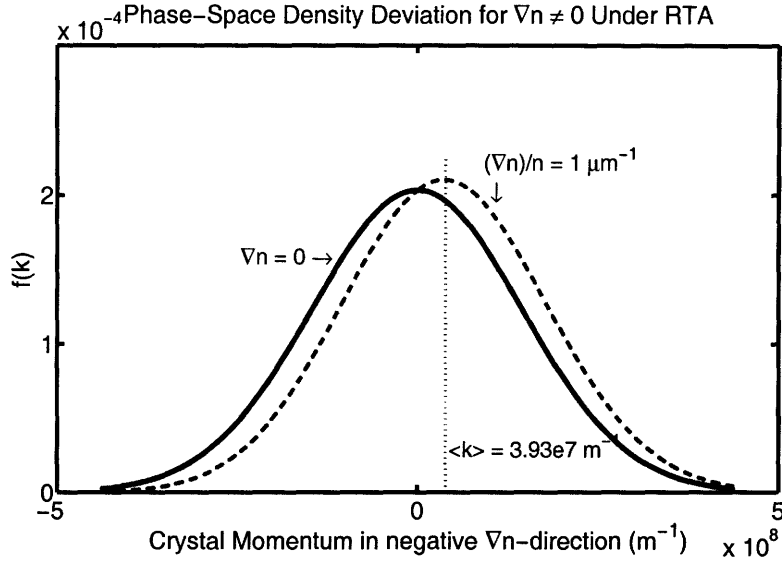


Figure 2-2: Plot of phase-space density as a function of  $\nabla n$ -directed wave-vector. The average wave-vector for the deviated distribution is  $\langle k \rangle = +3.93e7$ , indicating a flux of carriers to the right. The situation is identical to that in Figure 2-1 except that  $\vec{E} = 0$  and  $\nabla n \neq 0$ .

to which their occupancy is given is not ambiguous. Since every particle in such a mode is always at the same position, the rest of the system (carriers *and* lattice here) with which it is interacting through local scattering processes all have the same temperature, given by the specification of the solution we seek. In a sense, the distribution function  $f = f_0 + \delta f_{\nabla T}$  that we are solving for is the solution to the BTE under RTA which has the property that these unambiguous  $k_z = 0$  carriers are distributed according to the local temperature at all points in space.

For the case of a density gradient, the same is true. The distribution function  $f = f_0 + \delta f_{\nabla n}$  that we are solving for is the solution to the BTE under RTA which has the property that the unambiguous  $k_z = 0$  carriers are distributed with the proper electro-chemical potential  $E_F$  so that with the local bandstructure, the Fermi-Dirac distribution would give us the carrier occupancy as pre-specified.

Note, however, that this does not imply that we have no current. We just observed that each slice of our distribution comes from an equilibrium occupancy function, which if it were to be used to occupy all states at a given position, would yield no net current at that location. Nevertheless, when we look at a fixed point in space, we find

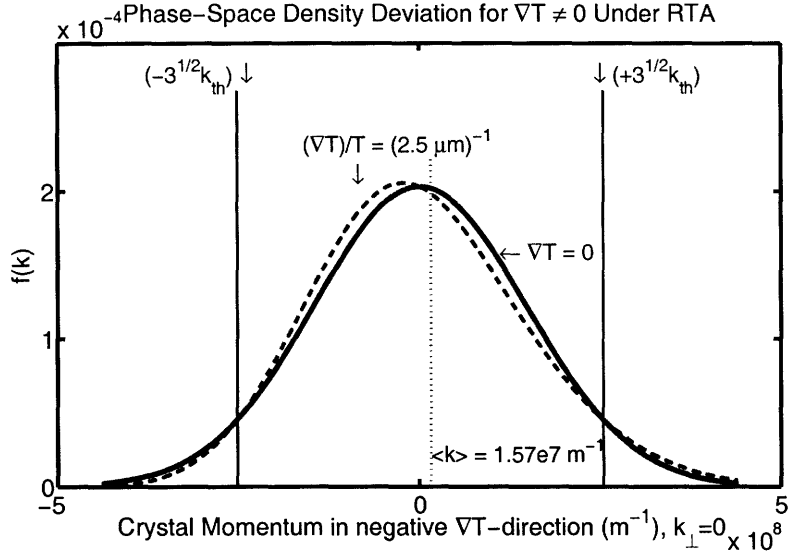


Figure 2-3: Plot of phase-space density as a function of  $\nabla T$ -directed wave-vector. Note that although the peak is moved to the left, the average wave-vector for the deviated distribution is  $\langle k \rangle = +1.57e7$ , indicating a flux of carriers to the right. The situation is identical to that in Figure 2-1 except that  $\vec{E} = 0$  and  $\nabla T \neq 0$ .

that having parts of the distribution originate at different locations (see Figure 2-4) leads to a net current. This is our picture of diffusion.

The complications of the proper definition above, however, can be put aside as we examine the picture that the RTA expression for  $f$  draws for us. In (2.3) we are being told that the value for the phase space density at some point in space is found using the ordinary Fermi-Dirac function, but using density and temperature from a different position than  $\vec{x}$ , while accessing the usual wave-vector  $\vec{k}$ . The position we are asked to access is to the left of the usual position for right-moving particles and is to the right of the usual position for left-moving particles. Moreover, the distance we are asked to shift by is precisely the mean-free-path for particles with momentum  $\vec{k}$ :  $\vec{l} = \vec{v}\tau = (\hbar\vec{k}/m^*)\tau(\vec{k})$ . Essentially, the picture of  $f(\vec{k})$  which (2.3) is drawing for us suggests that the states which compose the distribution at rightward momenta, whether filled with electrons or holes, on average originated one mean-free-path to the left, and so should be occupied as though part of a distribution with density and temperature as found at that location.

Recent efforts [24] have been made to engineer the energy-dependence of the re-

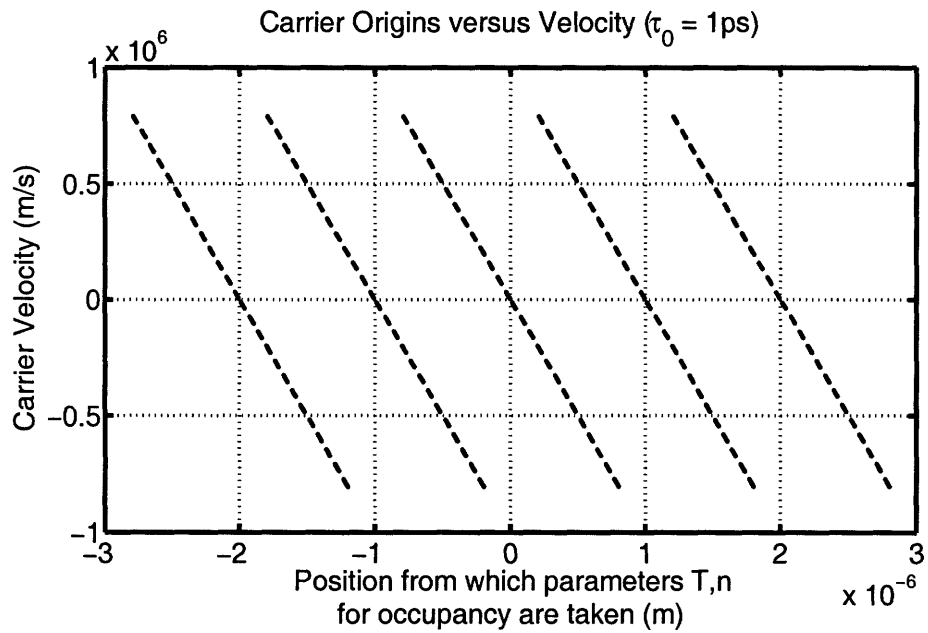


Figure 2-4: Plot of  $f_0$  arguments in (2.3) for  $\tau(\vec{k}) = \tau_0$ . Positive-velocity phase-space volume elements come from negative relative locations and negative-velocity elements come from positive relative locations. The positions whose actual occupancy these origination-locations inform can be seen by following any branch to zero momentum where the mean-free-path vanishes.

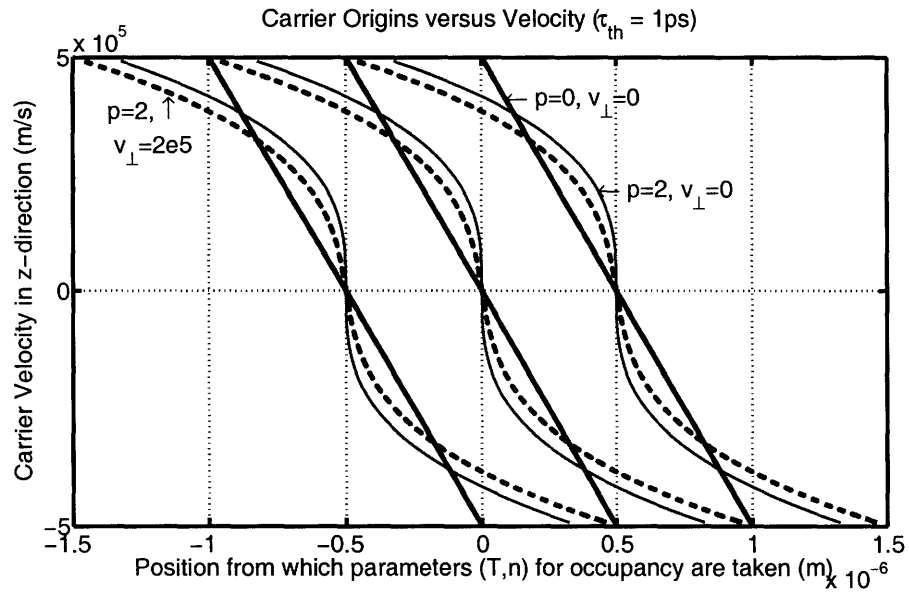


Figure 2-5: Plot of  $f_0$  arguments in (2.3) for  $\tau(\vec{k}) = \tau_{th} \left( \frac{|\vec{k}|}{\sqrt{2}k_{th}} \right)^p$ , where  $p$  is known as the scattering parameter. The general description of the  $p = 0$  lines are the same as in Figure 2-4. For  $\vec{k}$  along the plotted direction  $\hat{z}$ , we see that the origination shift follows a simple power-law in  $|\vec{k}|$ . For the  $\vec{k}_\perp \neq 0$  case, the behavior at small  $k_z$  resembles that of the  $p = 0$  case, while at large  $k_z$  the origination shift follows the  $\vec{k}_\perp = 0, p \neq 0$  result.

laxation time  $\tau$  to improve the Seebeck coefficient and thereby the thermoelectric performance [24, 2]. For this reason we have included the corresponding plot of origination location with velocity in the case of  $\tau(\vec{k})$  following a power-law with power  $p$  Figure 2-5, a situation for which we present analytical calculations of the transport coefficients in the subsequent sections. Not only is this case important because it is analytically tractable, but in principle, these coefficients can be approximated for arbitrary  $\tau$  in materials with spherical bandstructures. To see this, consider a short power-series expansion of  $\tau(|\vec{k}|)$  and interpret the calculated coefficients as resulting from parallel scattering channels.

## 2.4 Drift and Diffusion in the Boltzmann Limit

### 2.4.1 Field Driven Transport

Here we seek to calculate the response of the distribution  $f$  to a nonzero electric field. In particular, we wish to calculate the ratio of the particle flow induced to the field needed to produce it:  $\sigma/q$ . We begin with (2.2), assuming the ionized dopant density is low enough to put the semiconductor in the Boltzmann limit:

$$\begin{aligned}
 f_0(\vec{x}, \vec{k}) &= \frac{n_0}{(2\pi k_{th}^2)^{3/2}} \exp\left[\frac{-|\vec{k}|^2}{2k_{th}^2}\right] \quad \text{where } k_{th} \equiv \sqrt{\frac{m^* k_B T}{\hbar^2}} \\
 f &\approx f_0 + \frac{-q\tau(\vec{k})}{\hbar} \vec{E} \cdot \nabla_{\vec{k}} f_0 \\
 &= f_0 + \frac{-q\tau(\vec{k})}{\hbar} \vec{E} \cdot \frac{-\vec{k}}{k_{th}^2} f_0
 \end{aligned} \tag{2.4}$$

From here, taking  $\vec{E}$  to be in the  $\hat{z}$ -direction, we calculate the current:

$$\begin{aligned}
 \frac{J_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f \\
 &= qE_z \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \left( \frac{\tau(\vec{k}) k_z}{\hbar k_{th}^2} \right) f_0
 \end{aligned} \tag{2.5}$$

so that upon taking  $\tau(|\vec{k}|) \rightarrow \tau_0$ ,

$$\begin{aligned} \frac{J_z}{q} &= qE_z \left( \frac{\tau_0}{k_{th}^2 m^*} \right) \int_{\mathbb{B}} d^3\vec{k} k_z^2 f_0 \\ &= qE_z \left( \frac{\tau_0}{k_{th}^2 m^*} \frac{n_0}{(2\pi k_{th}^2)^{3/2}} \right) \int_{\mathbb{B}} d^3\vec{k} k_z^2 \exp \left[ \frac{-|\vec{k}|^2}{2 k_{th}^2} \right] \end{aligned} \quad (2.6)$$

A simple integral identity which we will find enormously useful in the case of  $\vec{k}$ -dependent  $\tau$  can be derived by a change of variables  $y = x^2$ :

$$\begin{aligned} I_l(\alpha) &\equiv \int_0^\infty dx x^l \exp[-\alpha x^2] \\ &= \frac{1}{2 \alpha^{\frac{(l+1)}{2}}} \Gamma \left( \frac{l+1}{2} \right) \end{aligned} \quad (2.7)$$

Since the Boltzmann limit causes the density of particles to drop exponentially at large wave-vectors, we may safely change the limits of integration from the zone-edge to  $\infty$  and apply the above result:

$$\begin{aligned} \frac{J_z}{q} &= qE_z \left( \frac{\tau_0}{k_{th}^2 m^*} \frac{n_0}{(2\pi k_{th}^2)^{3/2}} \right) \left( 2\pi \int_0^\pi d\theta \cos^2(\theta) \sin(\theta) \right) I_4 \left( \frac{1}{2 k_{th}^2} \right) \\ &= qE_z \left( \frac{\tau_0 n_0}{2^{3/2} \pi^{3/2} k_{th}^5 m^*} \right) \frac{4\pi}{3} \frac{1}{2} (2 k_{th}^2)^{5/2} \Gamma \left( \frac{5}{2} \right) \\ &= qE_z \left( \frac{4\tau_0 n_0}{3\sqrt{\pi} m^*} \right) \left( \frac{3\sqrt{\pi}}{4} \right) \\ &= qE_z \left( \frac{\tau_0 n_0}{m^*} \right) \implies \frac{\sigma}{q} = n_0 \frac{q\tau_0}{m^*} = \frac{\sigma_{\text{Drude}}}{q} \end{aligned} \quad (2.8)$$

Now we solve for the case of  $\tau(|\vec{k}|) = \tau_{th} \left( \frac{|\vec{k}|}{\sqrt{2} k_{th}} \right)^p$ . This power-law dependence of the relaxation time, with  $p$  the so-called “scattering parameter,” is of great interest in recent thermoelectric research [24], and as we will see in Chapter 3, finding the transport coefficients for arbitrary  $p \neq 0$  will permit us to match the analytical result for Seebeck coefficient from Section 1.2.3. Note that since both models assume a parabolic band-edge, we may relate the scattering parameter here to the one from



Section 1.2.3 with  $p = 2r$ . The calculation is nearly identical to the one above until:

$$\begin{aligned}
\frac{J_z}{q} &= qE_z \left( \frac{\tau_{th}}{\sqrt{2^p k_{th}^{2+p} m^*}} \frac{n_0}{(2\pi k_{th}^2)^{3/2}} \right) \left( 2\pi \int_0^\pi d\theta \cos^2(\theta) \sin(\theta) \right) I_{4+p} \left( \frac{1}{2 k_{th}^2} \right) \\
&= qE_z \left( \frac{\tau_{th} n_0}{2^{(3+p)/2} \pi^{3/2} k_{th}^{5+p} m^*} \right) \frac{4\pi}{3} \frac{1}{2} (2 k_{th}^2)^{(5+p)/2} \Gamma \left( \frac{5+p}{2} \right) \\
&= qE_z \left( \frac{4 \tau_{th} n_0}{3\sqrt{\pi} m^*} \right) \Gamma \left( \frac{5+p}{2} \right) \implies \frac{\sigma}{q} = n_0 \frac{q\tau_{th}}{m^*} \frac{\Gamma \left( \frac{5+p}{2} \right)}{\Gamma \left( \frac{5}{2} \right)}
\end{aligned} \tag{2.9}$$

The current driven by a given electric field is amplified by a ratio of gamma functions which clearly goes to unity as the scattering time becomes independent of wave-vector (or energy). We note that this result, though found by the author independently, is also found in Lundstrom's *Fundamentals of Carrier Transport* [23].

## 2.4.2 Density Gradient Driven Transport

For  $\nabla n \neq 0$ , in the Boltzmann limit, we may calculate the deviation in the phase-space density  $f$  and use it to derive a diffusion coefficient. Beginning with (2.3):

$$\begin{aligned}
f_0(\vec{x}, \vec{k}) &= \frac{n(\vec{x})}{(2\pi k_{th}^2)^{3/2}} \exp \left[ \frac{-|\vec{k}|^2}{2 k_{th}^2} \right] \quad \text{where } k_{th} \equiv \sqrt{\frac{m^* k_B T}{\hbar^2}} \\
f(\vec{x}_0, \vec{k}) &\approx f_0(\vec{x}_0, \vec{k}) + \frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot \left( \nabla f_0(\vec{x}, \vec{k}) \Big|_{\vec{x}=\vec{x}_0} \right) \\
&= f_0(\vec{x}_0, \vec{k}) \left( 1 + \frac{-\hbar\tau(\vec{k})}{m^*} \frac{\vec{k} \cdot (\nabla n)}{n(\vec{x}_0)} \right)
\end{aligned} \tag{2.10}$$

where we have chosen to specify a location  $\vec{x}_0$  at which to evaluate the current simply to avoid confusion. Since the solution we seek has a constant nonzero value of  $\nabla n$  across the device, we may choose to define the profile as  $n(\vec{x}) = n_0 + (\nabla n)(\vec{x} - \vec{x}_0)$ , so that the local density shares the  $n_0$  convention with the electric field case. As in the previous section, we again note that only the deviation  $\delta f_{\nabla n}$  contributes to the

current:

$$\begin{aligned}
\frac{J_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f \\
&= \frac{-\nabla n}{n_0} \int_{\mathbb{B}} d^3 \vec{k} \left( \frac{\hbar k_z}{m^*} \right) \left( \frac{\hbar \tau(\vec{k}) k_z}{m^*} \right) f_0 \\
&= (-\nabla n) \frac{(\hbar/m^*)^2}{(2\pi k_{th}^2)^{3/2}} \int_{\mathbb{B}} d^3 \vec{k} \tau(\vec{k}) k_z^2 \exp \left[ \frac{-|\vec{k}|^2}{2 k_{th}^2} \right]
\end{aligned} \tag{2.11}$$

Again taking  $\tau(|\vec{k}|) = \tau_0$ :

$$\begin{aligned}
\frac{J_z}{q} &= (-\nabla n) \frac{(\hbar/m^*)^2 \tau_0}{(2\pi k_{th}^2)^{3/2}} \left( 2\pi \int_0^\pi d\theta \cos^2(\theta) \sin(\theta) \right) I_4 \left( \frac{1}{2k_{th}^2} \right) \\
&= (-\nabla n) \frac{(\hbar/m^*)^2 \tau_0}{(2\pi k_{th}^2)^{3/2}} \left( \frac{4\pi}{3} \right) \frac{1}{2} (2k_{th}^2)^{5/2} \Gamma \left( \frac{5}{2} \right) \\
&= (-\nabla n) \left( \frac{\hbar k_{th}}{m^*} \right)^2 \tau_0 \left( \frac{4}{3\sqrt{\pi}} \right) \Gamma \left( \frac{5}{2} \right) \\
&= (-\nabla n) \frac{k_B T \tau_0}{m^*} \implies D = \frac{k_B T}{q} \frac{q \tau_0}{m^*} = \frac{k_B T}{q} \mu_{\text{Drude}}
\end{aligned} \tag{2.12}$$

so our solution satisfies the Einstein relation as expected. Now we turn to case where

$\tau(|\vec{k}|) = \tau_{th} \left( \frac{|\vec{k}|}{\sqrt{2} k_{th}} \right)^p$ :

$$\begin{aligned}
\frac{J_z}{q} &= (-\nabla n) \frac{(\hbar/m^*)^2 \tau_{th}}{(2k_{th}^2)^{p/2} (2\pi k_{th}^2)^{3/2}} \left( 2\pi \int_0^\pi d\theta \cos^2(\theta) \sin(\theta) \right) I_{4+p} \left( \frac{1}{2k_{th}^2} \right) \\
&= (-\nabla n) \frac{(\hbar/m^*)^2 \tau_{th}}{2^{(3+p)/2} k_{th}^{3+p} \pi^{3/2}} \left( \frac{4\pi}{3} \right) \frac{1}{2} (2k_{th}^2)^{(5+p)/2} \Gamma \left( \frac{5+p}{2} \right) \\
&= (-\nabla n) \left( \frac{\hbar k_{th}}{m^*} \right)^2 \tau_{th} \frac{\Gamma \left( \frac{5+p}{2} \right)}{\Gamma \left( \frac{5}{2} \right)} \implies D = \frac{k_B T}{q} \frac{q \tau_{th}}{m^*} \frac{\Gamma \left( \frac{5+p}{2} \right)}{\Gamma \left( \frac{5}{2} \right)}
\end{aligned} \tag{2.13}$$

which again is a simple modification of the diffusion coefficient for  $p = 0$  and  $\tau_{th} = \tau_0$ .

### 2.4.3 Temperature Gradient Driven Transport

We will now do the analogous calculation for the Soret coefficient, defined by

$$\frac{\vec{J}_z}{q} = S(-\nabla T) \quad \text{where} \quad \nabla n = \vec{E} = 0 \quad \text{and} \quad |\nabla T| = (\nabla T) \cdot \hat{z} \tag{2.14}$$

We note that this definition is totally different from the definition of the Seebeck coefficient found in many texts [3], which looks identical to this definition, except that the gradient of the electro-chemical potential satisfy  $\nabla E_F = 0$ . While in most semiconductor device problems, we presume that gradients in  $E_F$  arise from either density gradients or electric fields, we must remember that  $E_F$  is directly dependent on  $T$ , a point which has recently warranted explicit clarification [14, 25] in the thermoelectricity literature.

While the changes in the transport coefficients  $\sigma$  and  $D$  in materials with nonzero scattering parameter  $p$  had identical functional form, we will find that the case for the Soret coefficient differs slightly. This difference, as we will see in Chapter 3 accounts for the improved Seebeck with increased  $p$ . We will draw a physical picture explaining the origin of the difference after we have derived it algebraically, which should present an intuitive picture for why materials with  $p > 0$  may make good thermoelectrics.

The basic Soret coefficient calculation begins as before. We define  $T(\vec{x}) = T_0 + (\nabla T)(\vec{x} - \vec{x}_0)$ . Recalling that our definition for the thermal wave-vector included the temperature:

$$\begin{aligned}
f_0(\vec{x}, \vec{k}) &= \frac{n_0}{(2\pi k_{th}^2)^{3/2}} \exp\left[\frac{-|\vec{k}|^2}{2k_{th}^2}\right] = \frac{n_0 \hbar^3}{(2\pi m^* k_B T(\vec{x}))^{3/2}} \exp\left[\frac{-\hbar^2 |\vec{k}|^2}{2m^* k_B T(\vec{x})}\right] \\
f(\vec{x}_0, \vec{k}) &\approx f_0(\vec{x}_0, \vec{k}) + \frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot \left(\nabla f_0(\vec{x}, \vec{k})\Big|_{\vec{x}=\vec{x}_0}\right) \\
&= f_0(\vec{x}_0, \vec{k}) + \left(\frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot (\nabla T)\right) \left(\frac{\partial f_0}{\partial T}\Big|_{\vec{x}=\vec{x}_0}\right) \\
&= f_0(\vec{x}_0, \vec{k}) + \left(\frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot (\nabla T)\right) \left(\frac{-3}{2T_0} + \frac{\hbar^2 |\vec{k}|^2}{2m^* k_B T_0^2}\right) f_0(\vec{x}_0, \vec{k}) \\
&= f_0(\vec{x}_0, \vec{k}) + \left(\frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot (\nabla T)\right) \left(\frac{-3}{2T_0} + \frac{|\vec{k}|^2}{2k_{th}^2 T_0}\right) f_0(\vec{x}_0, \vec{k}) \\
&\qquad\qquad\qquad \text{where } k_{th} \equiv \sqrt{\frac{m^* k_B T_0}{\hbar^2}}
\end{aligned} \tag{2.15}$$

Moving forward as before, we now calculate the particle flow under temperature

gradient:

$$\begin{aligned}
\frac{J_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f \\
&= \frac{(-\nabla T)}{T_0} \int_{\mathbb{B}} d^3 \vec{k} \left( \frac{\hbar k_z}{m^*} \right) \left( \frac{\hbar \tau(\vec{k}) k_z}{m^*} \right) \left( \frac{-3}{2} + \frac{|\vec{k}|^2}{2k_{th}^2} \right) f_0 \\
&= \frac{(-\nabla T)}{T_0} \frac{(\hbar/m^*)^2 n_0}{(2\pi k_{th}^2)^{3/2}} \int_{\mathbb{B}} d^3 \vec{k} \tau(\vec{k}) k_z^2 \left( \frac{-3}{2} + \frac{|\vec{k}|^2}{2k_{th}^2} \right) \exp \left[ \frac{-|\vec{k}|^2}{2k_{th}^2} \right]
\end{aligned} \tag{2.16}$$

Again taking  $\tau(|\vec{k}|) = \tau_0$ :

$$\begin{aligned}
\frac{J_z}{q} &= \frac{(-\nabla T)}{T_0} \frac{(\hbar/m^*)^2 n_0 \tau_0}{(2\pi k_{th}^2)^{3/2}} \left( 2\pi \int_0^\pi d\theta \cos^2(\theta) \sin(\theta) \right) \\
&\quad \cdot \left[ \frac{-3}{2} I_4 \left( \frac{1}{2k_{th}^2} \right) + \frac{1}{2k_{th}^2} I_6 \left( \frac{1}{2k_{th}^2} \right) \right] \\
&= \frac{(-\nabla T)}{T_0} \frac{(\hbar/m^*)^2 n_0 \tau_0}{(2k_{th}^2)^{3/2}} \left( \frac{4}{3\sqrt{\pi}} \right) \left[ \frac{-3}{4} (2k_{th}^2)^{5/2} \Gamma \left( \frac{5}{2} \right) + \frac{1}{2} (2k_{th}^2)^{5/2} \Gamma \left( \frac{7}{2} \right) \right] \\
&= \frac{(-\nabla T)}{T_0} \left( \frac{\hbar k_{th}}{m^*} \right)^2 n_0 \tau_0 \left( \frac{4}{3\sqrt{\pi}} \right) \left[ \frac{-3}{2} \Gamma \left( \frac{5}{2} \right) + \Gamma \left( \frac{7}{2} \right) \right] \\
&= (-\nabla T) \left( \frac{k_B}{m^*} \right) n_0 \tau_0 \implies S = \frac{k_B}{q} n_0 \frac{q\tau_0}{m^*} = \frac{n_0 k_B}{q} \mu_{\text{Drude}}
\end{aligned} \tag{2.17}$$

And for  $\tau(|\vec{k}|) = \tau_{th} \left( \frac{|\vec{k}|}{\sqrt{2} k_{th}} \right)^p$ :

$$\begin{aligned}
\frac{J_z}{q} &= \frac{(-\nabla T)}{T_0} \frac{(\hbar/m^*)^2 n_0 \tau_{th}}{(2k_{th}^2)^{(3+p)/2}} \left( \frac{4}{3\sqrt{\pi}} \right) \\
&\quad \cdot \left[ \frac{-3}{2} I_{4+p} \left( \frac{1}{2k_{th}^2} \right) + \frac{1}{2k_{th}^2} I_{6+p} \left( \frac{1}{2k_{th}^2} \right) \right] \\
&= \frac{(-\nabla T)}{T_0} \left( \frac{\hbar k_{th}}{m^*} \right)^2 n_0 \tau_{th} \left( \frac{4}{3\sqrt{\pi}} \right) \\
&\quad \cdot \left[ \frac{-3}{2} \Gamma \left( \frac{5+p}{2} \right) + \Gamma \left( \frac{7+p}{2} \right) \right] \\
&= \frac{(-\nabla T)}{T_0} \left( \frac{\hbar k_{th}}{m^*} \right)^2 n_0 \tau_{th} \left( \frac{4}{3\sqrt{\pi}} \right) \left[ \frac{-3}{2} + \frac{5+p}{2} \right] \Gamma \left( \frac{5+p}{2} \right) \\
&= \frac{(-\nabla T)}{T_0} \left( \frac{k_B}{m^*} \right) n_0 \tau_{th} \left( 1 + \frac{p}{2} \right) \frac{\Gamma \left( \frac{5+p}{2} \right)}{\Gamma \left( \frac{5}{2} \right)} \implies S = \frac{k_B}{q} n_0 \frac{q\tau_{th}}{m^*} \frac{\Gamma \left( \frac{5+p}{2} \right)}{\Gamma \left( \frac{5}{2} \right)} \left( 1 + \frac{p}{2} \right)
\end{aligned} \tag{2.18}$$

which differs slightly from the modifications to  $D$  and  $\sigma$  for cases of nonzero  $p$ . In particular, for larger  $p$ , the ratio of the Soret coefficient to the diffusion coefficient and conductivity increases as  $(1 + p/2)$ .

To see why, we examine the dependence of the gaussian  $f_0(\vec{k})$  on the temperature  $T$ . The temperature appears twice in the definition, once in the exponent to widen the gaussian at higher temperatures, but again in the multiplicative factor in front to retain the normalization of  $\int_{\mathbb{B}} d^3\vec{k} f_0(\vec{k}) = n_0$ . The result is that at low wave-vectors, an increased temperature means the occupancy  $f$  is reduced, while at high wave-vectors, an increased temperature means  $f$  is increased. This combines with the wave-vector-based origination intuition from Figure 2-5 to create an unexpected situation in which low wave-vector carriers diffuse up a temperature gradient while high wave-vector carriers diffuse down it. This is the reason for the negative sign on the  $3/2$  in (2.15). In the end, however, in the Boltzmann regime where we have done our calculations to this point, this negative flow ( $\propto -3/2$ ) is overcome by the contribution to the current from large wave-vectors ( $\propto 5/2$  in (2.17) and  $\propto (5 + p)/2$  in (2.18)).

In fact, if we examine the situation in one dimension analytically, we see:

$$\frac{\partial}{\partial T} \left( \frac{n_0 \hbar^3}{(2\pi m^* k_B T)^{1/2}} \exp \left[ \frac{-\hbar^2 k^2}{2m^* k_B T} \right] \right) = \left( \frac{-1}{2} + \frac{k^2}{2k_{th}^2} \right) \frac{1}{T} f_0 = 0 \implies k = k_{th} \quad (2.19)$$

which we can interpret as quantifying the above relative statements about low and high wave-vectors. Note that while our definition of  $\tau(\vec{k})$  above was made to give  $\tau_{th}$  dimensions of time, our choice for the scale of the denominator is also convenient in that (up to a factor of  $O(1)$ , as seen in the plot Figure 2-5), for  $p > 0$ ,  $\tau < \tau_{th}$  for wave-vectors whose carriers contribute negatively to diffusion and  $\tau > \tau_{th}$  for wave-vectors whose carriers contribute positively to diffusion. These larger relaxation times for positively diffusing carriers increase the magnitude of the Soret coefficient, which in turn raises the Seebeck coefficient as we will see in Chapter 3, and the ultimately overall figure-of-merit  $ZT$ .

## 2.5 Outside the Boltzmann Limit

While the integrals are simpler to compute analytically in the Boltzmann limit, for most applications thermoelectric materials are doped so that their equilibrium  $E_F$  is near the bandedge where, as we will show in Section 3.5.1,  $ZT$  is maximized. In this case we can still numerically compute the quantities of interest. The validity of their application in the real-space formalism of Chapter 3 remains intact because heavily-doped semiconductors can still be described by the BTE and the relaxation-lengths are still sufficiently short to warrant use of the RTA. While their doping densities do not permit us to describe their occupancies in terms of simple gaussians as before, as long as we take care to distinguish between the actual impurity and the more easily measured ionized impurity concentration (denoted  $n_D$  for the n-type materials in this thesis), our method remains valid.

Outside the Boltzmann limit, the equilibrium phase-space density for single-valley spherical, parabolic bandstructures has the same  $\vec{x}$ - and  $\vec{k}$ -dependence as the Fermi-Dirac occupancy because the available states (per unit volume) are evenly-spaced in  $\vec{k}$ -space with two states per  $(2\pi)^3$ :

$$f_0(\vec{x}, \vec{k}) = \frac{2}{(2\pi)^3} \frac{1}{\exp\left[\frac{|\vec{k}|^2}{2k_{th}} + \frac{E_C - E_F}{k_B T}\right] + 1} \quad (2.20)$$

Since we wish to describe the occupancy in terms of  $(n, T, \phi)$  instead of  $(E_F, T)$ , we can integrate over the bandstructure to find the density  $n$  and substitute back using this relationship. Since parabolic bandstructure makes occupancy at high  $|\vec{k}|$  unlikely, we are still free to change our integrals over the Brillouin zone  $\mathbb{B}$  into integrals over

all of  $\vec{k}$ -space:

$$\begin{aligned}
n(\vec{x}) &= \int_{\mathbb{B}} d^3\vec{k} f_0(\vec{x}, \vec{k}) \\
&= 4\pi \int_0^\infty k^2 dk \frac{2}{(2\pi)^3} \frac{1}{\exp\left[\frac{|\vec{k}|^2}{2k_{th}} + \frac{E_C - E_F}{k_B T}\right] + 1} \\
&= \frac{1}{\pi^2} \int_0^\infty k^2 dk \frac{1}{\exp\left[\frac{|\vec{k}|^2}{2k_{th}}\right] A + 1} \quad \text{where } A \equiv \exp\left[\frac{E_C - E_F}{k_B T}\right] \quad (2.21) \\
&\implies A(\vec{x}) = A(n(\vec{x}), T(\vec{x})) \\
&\implies f_0(\vec{x}, \vec{k}) = \frac{1}{4\pi^3} \frac{1}{\exp\left[\frac{|\vec{k}|^2}{2k_{th}}\right] A(n, T) + 1}
\end{aligned}$$

Then to find the transport coefficients  $\sigma/q$ ,  $D$ , and  $S$ , we simply apply the procedure as before. We begin with the RTA expression for the leading-order deviation in  $f$ , then find the induced particle flow from the first moment of  $f$ , and divide by the source gradient to find the transport coefficient. For  $\sigma/q$ :

$$\begin{aligned}
\delta f_{\vec{E}} &= \frac{-q\tau(\vec{k})}{\hbar} \vec{E} \cdot \nabla_{\vec{k}} f_0 \\
\frac{\vec{J}_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f = \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \delta f_{\vec{E}} \quad (2.22) \\
\implies \frac{\sigma}{q} &= \frac{-q}{\hbar} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \tau(\vec{k}) \left( \frac{\partial}{\partial k_z} f_0 \right)
\end{aligned}$$

Likewise, for  $D$ :

$$\begin{aligned}
\delta f_{\nabla n} &= \frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot (\nabla n) \left( \frac{\partial}{\partial n} f_0 \right) \\
\frac{\vec{J}_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f = \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \delta f_{\nabla n} \quad (2.23) \\
\implies D &= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \tau(\vec{k}) k_z \left( \frac{\partial}{\partial n} f_0 \right)
\end{aligned}$$

and for  $S$ :

$$\begin{aligned}\delta f_{\nabla T} &= \frac{-\hbar\tau(\vec{k})}{m^*} \vec{k} \cdot (\nabla T) \left( \frac{\partial}{\partial T} f_0 \right) \\ \frac{\vec{J}_z}{q} &= \left\langle \frac{\hbar k_z}{m^*} \right\rangle_f = \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \delta f_{\nabla T} \\ \implies S &= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_z}{m^*} \right) \tau(\vec{k}) k_z \left( \frac{\partial}{\partial T} f_0 \right)\end{aligned}\tag{2.24}$$

keeping in mind that  $A$  in the expression for  $f_0$  above is a function of  $n$  and  $T$ .

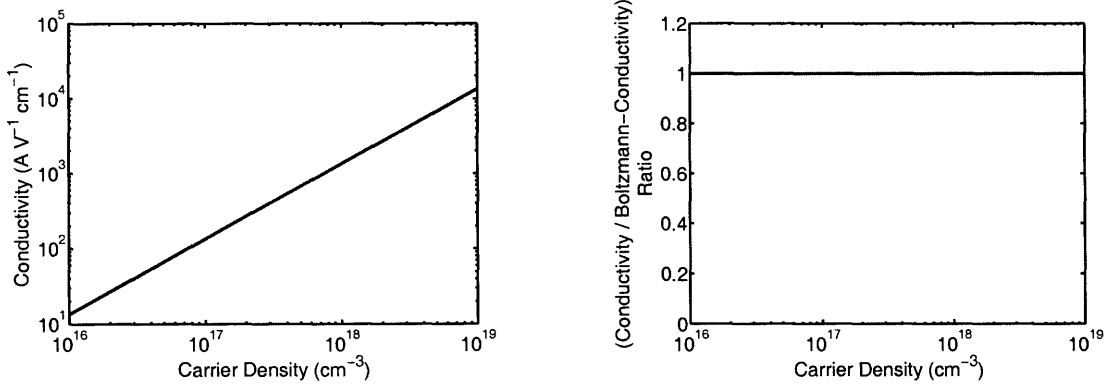


Figure 2-6: Left: Conductivity as a function of carrier density. Right: Ratio of conductivity to the value expected from extrapolating the Boltzmann-limit expression. Note that the relaxation time  $\tau$  is assumed here to be independent of  $n_D$ .

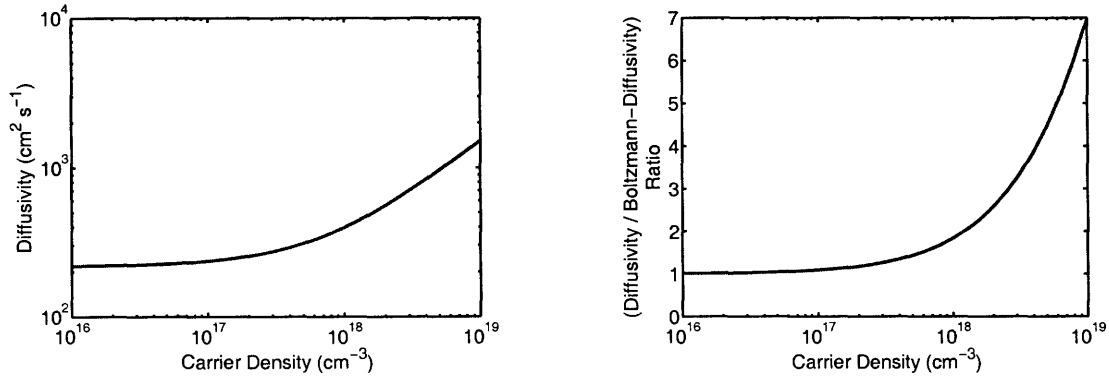


Figure 2-7: Left: Diffusivity as a function of carrier density. Right: Ratio of diffusivity to the value expected from extrapolating the Boltzmann-limit expression. Again, the relaxation time  $\tau$  is assumed here to be independent of  $n_D$ .

A numerical solution was developed in MATLAB, the source code for which appears in Section C.1. The primary purpose of this code is to produce transport



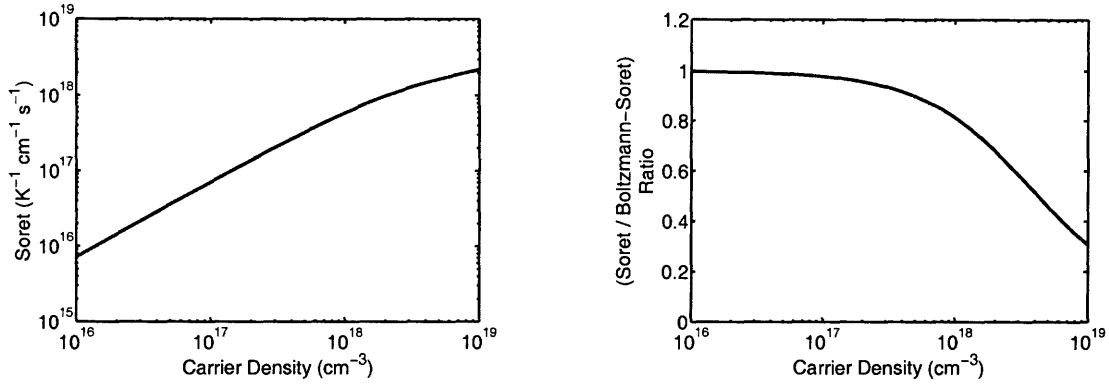


Figure 2-8: Left: Soret coefficient as a function of carrier density. Right: Ratio of Soret coefficient to the value expected from extrapolating the Boltzmann-limit expression. Again, the relaxation time  $\tau$  is assumed here to be independent of  $n_D$ .

coefficients to feed into the simulator discussed in Chapter 3, but here we plot the coefficients as a function of density. In order to convert from density to equilibrium Fermi level, we provide Figure 2-9.

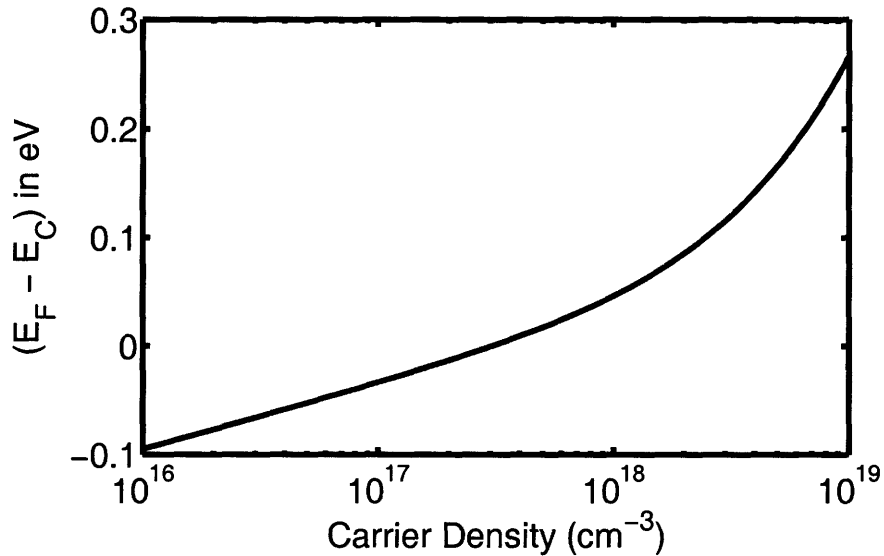


Figure 2-9:  $E_F$  versus  $n_D$  for n-GaAs. Spherical parabolic bandstructure assumed, with  $m^*$  taken as  $0.063m_0$ .

In Figure 2-6 we find that the Boltzmann-limit expression for the conductivity remains valid as the density increases and the equilibrium Fermi level rises into the band. Intuitively, we may expect the conductivity expression to remain unchanged

because each of the carriers still absorbs the same momentum from a given electric field in a relaxation time  $\tau$ .

In Figure 2-7 we find that the Boltzmann-limit expression for the diffusivity underestimates the diffusivity as the density increases. Though a derivation of the analytical expression for the diffusivity in the degenerate limit [26] would take us too far afield, considering this case provides some clarity. In the Boltzmann limit, the diffusivity can be seen as independent of density by the Einstein relation  $D = \frac{k_B T}{q} \mu$ . By contrast, in the degenerate limit for a density gradient  $\nabla n \propto -\hat{z}$ , the  $+k_z$  part of the Fermi surface is occupied as though part of a distribution with greater density than the  $-k_z$  part. However, the difference in  $E_F$  that this corresponds to is dependent on the density of states at the Fermi level  $D(E_F)$ . It should therefore come as no surprise that in the degenerate limit, the diffusivity and conductivity obey a different Einstein relation where  $\sigma/D \propto D(E_F)$ . Since  $E_F \propto n^{2/3}$ , and  $D(E) \propto \sqrt{E}$  in the parabolic band, we expect that  $D \propto n^1/n^{1/3} = n^{2/3}$ , so that the diffusivity should rise with increasing density, but not as quickly as the conductivity.

In Figure 2-8 we find that the Boltzmann-limit expression for the Soret coefficient over-estimates the Soret coefficient as the density increases. Using a similar picture to the diffusivity case, we note that for  $\nabla n \propto -\hat{z}$ , the  $+k_z$  part of the Fermi surface is occupied as though part of a distribution with higher temperature than the  $-k_z$  part. If the density of states  $D(E)$  is an increasing function of  $E$  at  $E_F$ , then the higher-temperature Fermi-Dirac occupancy from the same  $E_F$  leads to a greater density. Thus we expect more carriers near the  $+k_z$  part of the Fermi surface than the  $-k_z$  part, leading to a flow of carriers  $\propto D'(E)|_{E=E_F}$ . For the parabolic bandstructure assumed, in the degenerate limit  $D'(E) \propto E^{-1/2}$  so that we expect the Soret coefficient to eventually decrease with density as  $S \propto n^{-1/3}$ .

If increasing carrier density beyond the Boltzmann-limit increases the conductivity and diffusivity, but eventually temperature-driven diffusion becomes weak, we may expect that the thermoelectric effects may vanish in the degenerate limit. However, as we noted in Section 1.2.1, practical devices benefit from high electrical conductivity. Therefore we may expect that some value for the carrier concentration

may prove optimal for thermoelectric devices. Our intuition will be confirmed in Chapter 3.

## 2.6 Summary and Conclusions

In this chapter, we have developed a picture of the microscopic transport processes relevant to thermoelectricity in reciprocal-space. We began with the BTE under RTA and derived general expressions for leading-order deviations in the distribution function in the presence of gradients ( $\nabla\phi$ ,  $\nabla n$ , and  $\nabla T$ ). We then used these expressions to analytically derive their transport coefficients ( $\sigma$ ,  $D$ , and  $S$  respectively), for both energy-independent  $\tau$  and  $\tau(E) \propto E^r$ . We then concluded by presenting the general formulation of these transport coefficients valid outside the dilute Boltzmann limit where practical thermoelectric elements for power-conversion fall.

The purpose of the chapter was two-fold. First, we looked to develop a visual picture of the distribution function in these transport processes, backed by analytical rigor. We did so both to provide intuition and to make vivid the fact underlying all thermoelectric effects, that the thermal and electrical responses are linked by a common distribution of charge carriers. Secondly, our analytical expressions will serve to extend the model for the Seebeck effect in Chapter 3 to realistic dopant densities and to include the energy-dependence of relaxation times, a recent focus of thermoelectric material design [24].



# Chapter 3

## Real-Space Picture of Thermoelectric Transport

### 3.1 Introduction and Overview

The Seebeck coefficient is typically calculated by the method in Section 1.2.3, which relates an applied temperature difference  $\Delta T$  to the resulting difference in electrochemical potential  $\Delta E_F$ . Although the previous calculations made no specific claims about the origin of  $\Delta E_F$ , this difference is sometimes intuitively ascribed to an electric field and sometimes ascribed to a difference in carrier density. Since both interpretations carry the same correct sign, as diffusion and Coulomb repulsion both drive carriers away from regions of high carrier concentration, it is easy to forget that each of these pictures (see Figure 3-1) alone are incomplete.

In the first picture,  $\Delta E_F$  is explained as the consequence of the difference in carrier density giving rise to a difference in chemical potential across the device. For a gas of charge-neutral particles this description is valid, but in a thermoelectric element where the carriers are charged, Poisson's equation requires that the buildup in carrier density on the cold-side and the depletion of carriers on the hot-side give rise to an electric field with implications for the flux of charged carriers in the center. This electric field and its associated contribution to the particle-flux balance is omitted from the story. Meanwhile, since the device is net charge neutral, Poisson's equation

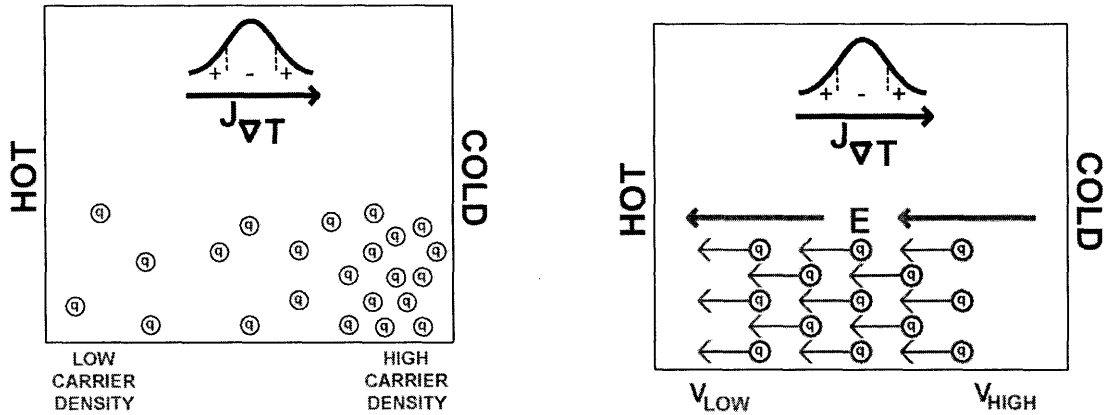


Figure 3-1: Visual depiction of Seebeck effect. Left: density-driven picture. Right: field-driven picture.

tells us there is negligible electric field flux at the ends. Since the density profile was chosen to everywhere cancel the temperature-driven diffusion, we have calculated the net flux to be nonzero in middle of the device yet zero at the ends, which is inconsistent with the steady-state continuity equation.

In the second picture,  $\Delta E_F$  is explained as the consequence of an electric field within the material which drives a current of carriers opposite to the current due to diffusion from the applied temperature gradient. This description alone offers no charge source for the electric field, required by Poisson's equation.

In addition to the inconsistencies with Maxwell's Equations and boundary conditions, these pictures also neglect the important direct dependence of the chemical potential on the temperature [25, 14]. As we will see in Section 3.4, an important part of why we see a difference in  $E_F$  across a thermoelement, particularly so for lightly-doped semiconductors with large  $\alpha$ , is that to maintain a constant carrier density in the presence of a temperature difference, the  $E_F$ -bandedge separation must change as well.

In Section 3.2 we address the aforementioned concerns and develop a viable and quantitative picture of the Seebeck effect, consistent with both the laws governing carrier transport in semiconductors and Maxwell's Equations. In Section 3.3 we make use of the results from Chapter 2 to extend this analysis to include energy-

dependent scattering times. In Section 3.4 we examine the results of the preceding sections to decompose the Seebeck coefficient into three distinct components. In Section 3.5 we again make use of the results from Chapter 2 to further extend this analysis to include relevant doping densities and briefly address the issue of the more complex bandstructure found in real thermoelectric materials. Finally, in Section 3.6 we demonstrate that the closely-related Peltier effect can similarly be described in a microscopic real-space picture, identifying  $\vec{J} \cdot \vec{E}$  as not just Joule heating alone but the sum of Joule and Peltier heating.

## 3.2 The Microscopic Seebeck Effect

We simulate the Seebeck effect within a drift-diffusion framework for an extrinsic semiconductor doped in the Boltzmann limit, in a system presumed to have an energy-independent momentum relaxation-time  $\tau_0$  and a fixed thermal conductivity  $\kappa$ . The fixed- $\kappa$  assumption may be thought of as an assumption that  $\kappa_{ph} \gg \kappa_e$ , or alternatively that the deviations in the distribution of carriers are fractionally small so that their thermal conductivity  $\kappa_e$  is everywhere roughly equal, a fact which will be confirmed by our solution.

Although in Chapter 2 we derived expressions for the microscopic transport coefficients  $\sigma$ ,  $D$ , and  $S$  which were valid in systems with energy-dependent relaxation times  $\tau(|\vec{k}|)$  and outside the Boltzmann limit, for simplicity we will not make use of those general expressions until Section 3.3 and Section 3.5.1 respectively.

We begin by taking the first moment of the BTE to generate the Momentum Balance Equation (MBE), writing all quantities in terms of the density  $n$ , temperature  $T$ , and electric potential  $\phi$ , and solving it self-consistently with Poisson's equation under the boundary conditions appropriate for the Seebeck effect. We thereby not only calculate the Seebeck coefficient  $\alpha$ , but also generate a real-space picture of all quantities of interest. They form an intuitive picture which resolves the fundamental issues raised in Figure 3-1.

### 3.2.1 The Balance Equations

The BTE, which describes the evolution of the entire distribution function, can be used to generate a series of equations describing the evolution of particular moments of the distribution function without approximation. The resulting equations are known as the balance equations. The first of these equations is the intuitive continuity equation which expresses charge conservation. The second of these equations, the momentum balance equation, can be approximated to recover the well-known drift and diffusion equations familiar to device engineers.

The procedure for generating a balance equation is to take a moment of the entire BTE. Conceptually the BTE, which constrains the time-evolution of  $f$  based on the laws of classical mechanics, expresses what must happen to this probability density at each point  $(\vec{x}, \vec{k}, t)$ . Since the BTE must hold at each point in phase space, the weighted sum of such points represented by a particular moment of the BTE's left-hand side must still be equal to the identically-weighted sum of the equation's right-hand side. Thus, the balance equations we are about to derive may be regarded as conditions on any distribution function required by the BTE.

#### The Continuity Balance Equation

For the continuity equation, we take the  $k^0$  moment of the full BTE:

$$\int_{\mathbb{B}} d^3\vec{k} k^0 \left[ \left( \frac{\partial}{\partial t} + \frac{\hbar}{m^*} \vec{k} \cdot \nabla + \frac{q}{\hbar} \vec{E} \cdot \nabla_k \right) f(\vec{x}, \vec{k}, t) \right] = \int_{\mathbb{B}} d^3\vec{k} k^0 \left[ \hat{S} f(\vec{x}, \vec{k}, t) \right] \quad (3.1)$$

$$\frac{\partial}{\partial t} n(\vec{x}, t) - \frac{1}{q} \nabla \cdot \vec{J}(\vec{x}, t) = G(\vec{x}, t) - R(\vec{x}, t)$$

Where we have defined the generation rate  $G$  and recombination rate  $R$  in terms of the point-to-point scattering rates  $S$  (which are in principle calculable from Fermi's Golden Rule) as:

$$G(\vec{x}, t) \equiv \int_{\mathbb{B}} d^3\vec{k} \int_{\mathbb{B}} d^3\vec{k}' f(\vec{x}, \vec{k}', t) S(\vec{k}', \vec{k}) (1 - f(\vec{x}, \vec{k}, t))$$

$$R(\vec{x}, t) \equiv \int_{\mathbb{B}} d^3\vec{k} \int_{\mathbb{B}} d^3\vec{k}' f(\vec{x}, \vec{k}, t) S(\vec{k}, \vec{k}') (1 - f(\vec{x}, \vec{k}', t)) \quad (3.2)$$



The resulting equation is known as the continuity equation and expresses the conservation law for density, the  $k^0$  moment of the distribution. Note that we performed integration by parts above and set the surface term to zero on the assumption that our bandstructure and carrier concentration do not permit population of levels near the zone-edge.

### The Momentum Balance Equation

To explain the Seebeck coefficient, we must examine the second balance equation, found by taking the  $k^1$  moment of the BTE:

$$\int_{\mathbb{B}} d^3\vec{k} \frac{\hbar\vec{k}}{m^*} \left[ \left( \frac{\partial}{\partial t} + \frac{\hbar}{m^*} \vec{k} \cdot \nabla + \frac{q}{\hbar} \vec{E} \cdot \nabla_k \right) f(\vec{x}, \vec{k}, t) \right] = \int_{\mathbb{B}} d^3\vec{k} \frac{\hbar\vec{k}}{m^*} \left[ \hat{S}f(\vec{x}, \vec{k}, t) \right] \quad (3.3)$$

Examining each term independently and using summation notation where not explicit:

$$\begin{aligned} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} \frac{\partial}{\partial t} f(\vec{x}, \vec{k}, t) \right) &= \frac{\partial}{\partial t} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} f(\vec{x}, \vec{k}, t) \right) \\ &= \frac{\partial}{\partial t} \left( \frac{J_j}{q} \right) \end{aligned} \quad (3.4)$$

$$\begin{aligned} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} \left[ \frac{\hbar\vec{k}}{m^*} \cdot \nabla f(\vec{x}, \vec{k}, t) \right] \right) &= \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar k_j}{m^*} \left[ \frac{\hbar k_i}{m^*} \cdot \frac{\partial}{\partial x_i} f(\vec{x}, \vec{k}, t) \right] \right) \\ &= \frac{2}{m^*} \frac{\partial}{\partial x_i} \left[ \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2 k_j k_i}{2m^*} \right) f(\vec{x}, \vec{k}, t) \right] \\ &= \frac{2}{m^*} \sum_i \frac{\partial}{\partial x_i} W_{ji} \end{aligned} \quad (3.5)$$

$$\begin{aligned} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} \left[ \frac{q\vec{E}}{\hbar} \cdot \nabla_k f(\vec{x}, \vec{k}, t) \right] \right) &= \frac{-q}{m^*} \int_{\mathbb{B}} d^3\vec{k} f(\vec{x}, \vec{k}, t) \frac{\partial}{\partial k_i} (k_j E_i) \\ &= \frac{-qE_j}{m^*} n(\vec{x}, t) \end{aligned} \quad (3.6)$$

$$\begin{aligned}
& \int_{\mathbb{B}} d^3 \vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} \left[ \hat{S}f(\vec{x}, \vec{k}, t) \right] \right) \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \int_{\mathbb{B}} d^3 \vec{k}' (k_j [f(k')S(k', k)(1 - f(k)) - f(k)S(k, k')(1 - f(k'))]) \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \int_{\mathbb{B}} d^3 \vec{k}' (k'_j - k_j) [f(k)S(k, k')(1 - f(k'))] \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \int_{\mathbb{B}} d^3 \vec{k}' (-k_j) \left[ f(k)S(k, k')(1 - f(k')) \left( 1 - \frac{k'_j}{k_j} \right) \right] \\
&= \frac{-\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \left[ f(k)k_j \frac{1}{\tau_m(k)} \right] \\
&= \frac{-J_j}{q} \left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k
\end{aligned} \tag{3.7}$$

where we have made the following definitions:

$$\begin{aligned}
\frac{1}{\tau_m(k)} &\equiv \int_{\mathbb{B}} d^3 \vec{k}' \left[ S(k, k')(1 - f(k')) \left( 1 - \frac{k'_j}{k_j} \right) \right] \\
\left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k &\equiv \frac{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*} \frac{1}{\tau_m(k)}}{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*}}
\end{aligned} \tag{3.8}$$

While these definitions may seem merely convenient, they occlude the important distinction between the various relaxation times in the system, a point which will be directly addressed in Section 3.2.3. Combining these results, we arrive at the Momentum Balance Equation:

$$\frac{\partial}{\partial t} \vec{J} + \frac{2}{m^*} \nabla \overline{\overline{W}} - \frac{q \vec{E}}{m^*} n = -\frac{\vec{J}}{q} \left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k \tag{3.9}$$

which in steady-state simplifies to:

$$\frac{2\mu}{q} \nabla \overline{\overline{W}} - n\mu \vec{E} + \frac{\vec{J}}{q} = 0 \tag{3.10}$$

where the mobility is defined as:

$$\mu \equiv \frac{q}{m^*} \left( \left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k \right)^{-1} \tag{3.11}$$

We pause here to take note of what the balance equations are capable of telling us. The first balance equation, the particle continuity equation, describes the evolution of the carrier density  $n$  in terms of a current  $\vec{J}$ . The second balance equation, the MBE, describes for us the evolution of the current density  $\vec{J}$  in terms of the energy density  $\overline{\overline{W}}$ . We note that at each stage, the object whose conservation law is described by any particular balance equation is dependent on the next higher object in the distribution moment hierarchy. In this way, any finite number of balance equations is unresolved and cannot be completely solved by itself. In order to perform an actual calculation we must at some point close the hierarchy by stopping at one equation and approximating the next moment.

In the case of deriving the Seebeck coefficient, we choose to solve only the continuity equation and MBE and to approximate the energy density  $\overline{\overline{W}}$ . This approximation is justified by equations (2.4), (2.10), and (2.15), which tell us that the deviations ( $\delta f_{\vec{E}}$ ,  $\delta f_{\nabla n}$ , and  $\delta f_{\nabla T}$ ) due to the gradients in ( $\phi$ ,  $n$ , and  $T$ ) are odd in the wave-vector component parallel to each gradient, so do not contribute to a change in the even moment  $\overline{\overline{W}} \propto \langle |\vec{k}|^2 \rangle$ . Note that this argument applies to all elements of the energy-density tensor, not just the diagonal ones, and likewise informs us that only the diagonal elements of  $\overline{\overline{W}}$  will survive. The approximate calculation of  $\overline{\overline{W}}$  gives:

$$\begin{aligned}
\overline{\overline{W}}_{ii} &= \frac{\hbar^2}{2m^*} \langle k_i k_i \rangle_{f_0 + \delta f} \\
&= \frac{\hbar^2}{2m^*} \langle k_i k_i \rangle_{f_0} \\
&= \frac{\hbar^2}{2m^*} \int_{\mathbb{B}} d^3 \vec{k} \left( (k_i)^2 \frac{n}{(2\pi k_{th}^2)^{3/2}} \exp \left[ \frac{-|\vec{k}|^2}{2k_{th}^2} \right] \right) \\
&= \frac{\hbar^2}{2m^*} \frac{n}{(2\pi k_{th}^2)^{3/2}} 2\pi \int_0^\pi d\theta (\cos^2 \theta \sin \theta) \int_0^\infty dk \left( k^4 \exp \left[ \frac{-|\vec{k}|^2}{2k_{th}^2} \right] \right) \quad (3.12) \\
&= \frac{\hbar^2}{2m^*} \frac{n}{(2\pi k_{th}^2)^{3/2}} \frac{4\pi}{3} \left[ \frac{1}{2} (2k_{th}^2)^{5/2} \Gamma \left( \frac{5}{2} \right) \right] \\
&= \frac{\hbar^2 k_{th}^2}{m^*} n \frac{1}{2} \\
&= \frac{1}{2} n k_B T
\end{aligned}$$

again making use of our previous identity (2.7). By symmetry, we have recovered the solution for any dilute gas in equilibrium:

$$\overline{\overline{W}} = \begin{bmatrix} \frac{1}{2}nk_B T & 0 & 0 \\ 0 & \frac{1}{2}nk_B T & 0 \\ 0 & 0 & \frac{1}{2}nk_B T \end{bmatrix} \quad (3.13)$$

The boundary conditions applicable to the Seebeck effect dictate the direction of the temperature gradient, which in turn dictates the direction of the density and electric potential gradients, thereby permitting the relevant form of the MBE to be reduced to a single dimension. Moreover, the effect calls for electrical open-circuit boundary conditions, which make  $\vec{J}/q \rightarrow 0$ . Substituting our above expression for  $\overline{\overline{W}}$  from (3.13) into (3.10) and applying these observations about the boundary conditions leads us to a simple version of the steady-state Momentum Balance Equation:

$$\frac{2\mu}{q} \left( \frac{1}{2}k_B T \right) \left( -\frac{\partial n}{\partial z} \right) + \frac{2\mu}{q} \left( \frac{1}{2}nk_B \right) \left( -\frac{\partial T}{\partial z} \right) + n\mu E_z = 0 \quad (3.14)$$

or

$$D \left( -\frac{\partial n}{\partial z} \right) + S \left( -\frac{\partial T}{\partial z} \right) + \frac{\sigma}{q} E_z = 0 \quad (3.15)$$

where  $D$  is the diffusivity,  $S$  is the Soret coefficient, and  $\sigma$  is the electrical conductivity. Note that we have just re-calculated  $D$  and  $S$ , and found that they do indeed reconcile with the results from Chapter 2. Later, in Section 3.3 and Section 3.5.1, we will regard this part of the derivation as approximating the more careful results from Chapter 2.

The equation for carrier transport found in (3.14) is essentially the one-dimensional, steady-state, electrical open-circuit drift-diffusion equation, generalized to include diffusion due to temperature gradients at constant  $n$  and  $\phi$ . Note that the Einstein relation can be simply read off when  $\nabla T \rightarrow 0$ .

Before we move on to the proper self-consistent solution, we pause to note that (3.14) on its own reveals the relationship between the two distinct pictures in Figure 3-1. The two types of flux, traditional field-driven current and density-driven diffusion, must together combine to offset any flux driven by an applied temperature gradient

at every point in space. How they cooperate to share those duties while satisfying Maxwell's Equations is the focus of the next section.

### 3.2.2 Self-Consistent Solution

The generalized drift-diffusion equation (3.14) must be solved self-consistently with Maxwell's Equations. For the Seebeck effect, a steady-state solution of a system with electrical open-circuit boundary conditions, the relevant equation is Poisson's equation:

$$\vec{E}(z) = \frac{\vec{D}(z)}{\epsilon_r} = \frac{1}{\epsilon_r} \int_{-\infty}^z dz' \frac{\rho_{\text{free}}(z')}{\epsilon_0} \quad (3.16)$$

where  $\epsilon_r$  is the DC relative permittivity of the intrinsic semiconductor. By identifying the free charge density with the excess density of mobile carriers (with charge  $q$ ) over the density of ionized dopant atoms, and using the net charge neutrality of the extrinsic semiconductor to justify the standard boundary condition of  $\vec{E} \rightarrow 0$  for  $z \rightarrow \pm\infty$ , we arrive at:

$$\vec{E}(z) = \int_0^z dz' \frac{q(n(z') - n_D(z'))}{\epsilon_0 \epsilon_r} \quad (3.17)$$

where we take the length of semiconductor to run from 0 to  $L$ .

Note that both (3.14) and (3.17) will clearly permit the self-consistent solution of inhomogeneously-doped elements, where  $n_D = n_D(z)$ .

Now that we have expressed all of the equations which describe the relevant physics in terms of a common set of quantities, we want to ensure that the equations and unknown quantities represent a uniquely-solvable system. The drift-diffusion equation looks to describe the system in terms of the density  $n$ , temperature  $T$ , and electric field  $\vec{E}$ . Note that the transport coefficients  $D$ ,  $S$ , and  $\sigma/q$  are completely determined by the aforementioned quantities, even without the simplifying assumptions used to arrive at (3.14). If we choose to discretize space into  $N$  points, then, we begin with  $3N$  unknown quantities. In the Seebeck effect, our semiconductor is placed in isothermal contact with a pair of reservoirs, one hot at  $z = 0$  and one

cold at  $z = L$ , which combines with our fixed thermal conductivity assumption to provide a complete temperature profile  $T(z)$ . Removing these  $N$  quantities from our list, we are left to solve for the density and field at each of the  $N$  points in space. Since Poisson's equation and the relevant electrical boundary conditions, succinctly expressed in (3.17), uniquely determine the electric field for any given distribution of free carriers  $n(z)$ , its inclusion reduces the problem to solving for the  $N$  values of the density  $n$  at each point in space in a way that satisfies (3.14) at each point in space. Since (3.14) is essentially asking us to balance the flux due to gradients in the three sources dissected in Chapter 2, we must have well-defined gradients at the points where we ask the flux to be balanced. If we choose the fluxes to balance at the positions halfway between each of the  $N$  discrete points where  $n$  is defined, we have  $(N - 1)$  flux-balance equations and  $N$  unknown densities. We need only include the overall charge neutrality of the element, which is not yet expressed in (3.17), to find that we have  $N$  independent equations and  $N$  unknowns.

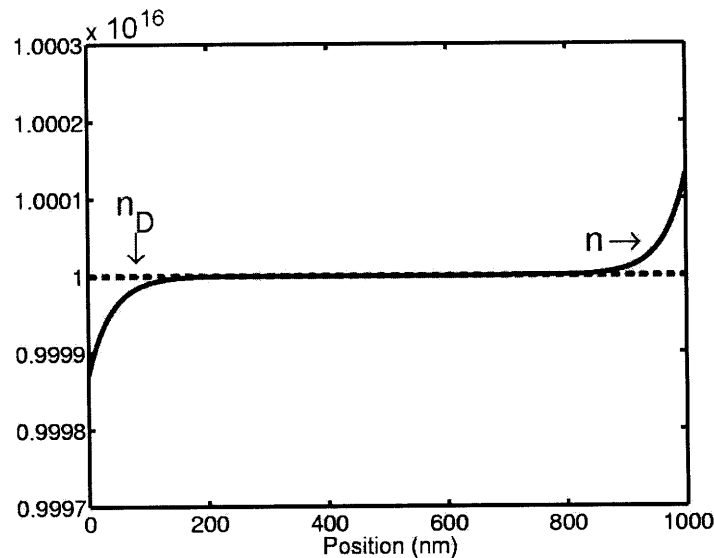


Figure 3-2: The carrier density profile for Seebeck effect at  $T_{\text{avg}} = 300\text{K}$ , shown with  $\Delta T = 1\text{K}$ . There is a buildup of free charge carriers on the cold-side (right) and a depletion on the hot-side (left). Parameters for the simulation taken from n-GaAs doped  $n_D = 1\text{e}16 \text{ cm}^{-3}$ :  $m^* = 0.063m_e$ ,  $\epsilon_r = 10.6$ . For reasons explained in the text, we assume a fixed value for thermal conductivity  $\kappa$  and take energy-independent momentum-relaxation time  $\tau_m = 0.3\text{ps}$ .

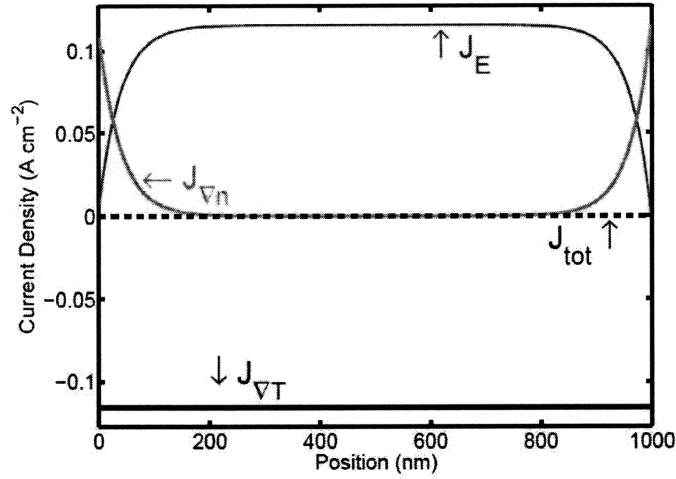


Figure 3-3: The source-specified currents sum to satisfy net flux balance at each point in space, as required by the steady-state electrical open-circuit boundary conditions in the Seebeck effect. All parameters for this simulation are identical to that from Figure 3-2.

To numerically compute the self-consistent solution of the generalized drift-diffusion equation (3.14) and Poisson's equation (3.17), we developed a MATLAB code (see Section C.2). The result is an intuitive real-space picture of the Seebeck effect, communicated in Figure 3-2, Figure 3-3, and Figure 3-4.

Upon examining the point-wise particle flux balance conveyed in Figure 3-3, we immediately see the proper interpretation of the relationship between the two intuitive pictures of the Seebeck effect from Figure 3-1. As seen in Figure 3-3, at each point in the device, the  $\nabla T$ -driven diffusion current must be counteracted by the combination of  $\nabla n$ -driven diffusion current and  $\vec{E}$ -driven drift current. At the edges of the device, where the overall charge neutrality of the device requires that  $\vec{E} \rightarrow 0$ , this duty is taken up by the density gradient. In addition to providing the current to oppose the applied  $\nabla T$ , this density gradient also serves as a layer (see Figure 3-5) of depleted carriers on the hot side and a layer of built-up excess carriers on the cold side. This charge configuration creates an electrostatic field which drives a drift current which takes over the duties of the density gradient for the interior of the device. Together, this picture provides a full semi-classical picture of the steady-state Seebeck effect in

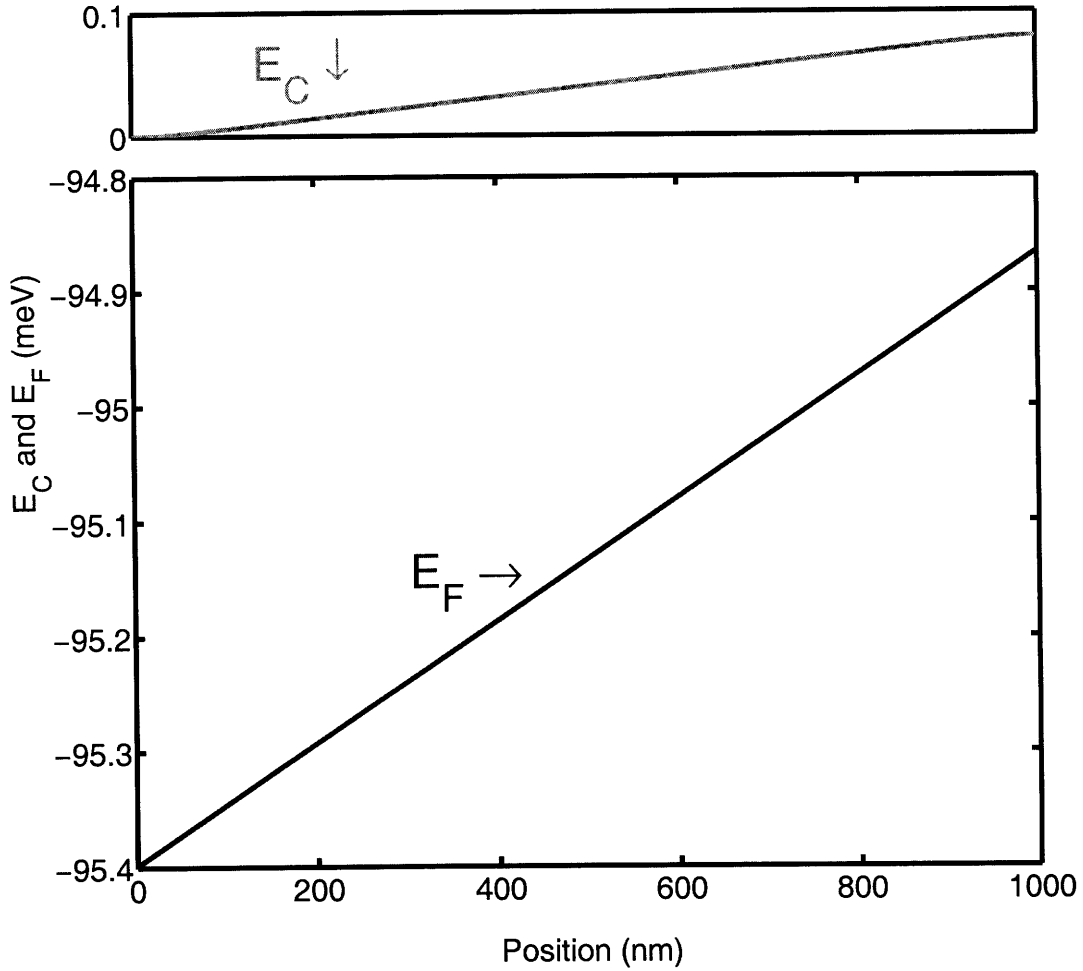


Figure 3-4: The band diagram for the Seebeck effect demonstrates that in the Boltzmann limit, only a fraction of the thermoelectric voltage can be ascribed to a gradient in the band-edge, and therefore gradient in the vacuum electric potential, through the bulk of the device. The  $86 \mu\text{V}/\text{K}$  limit explained in Section 3.4 can be seen, as  $\Delta T = 1\text{K}$  and  $\Delta E_C \approx 86 \mu\text{V}$ . All parameters for this simulation are identical to that from Figure 3-2.



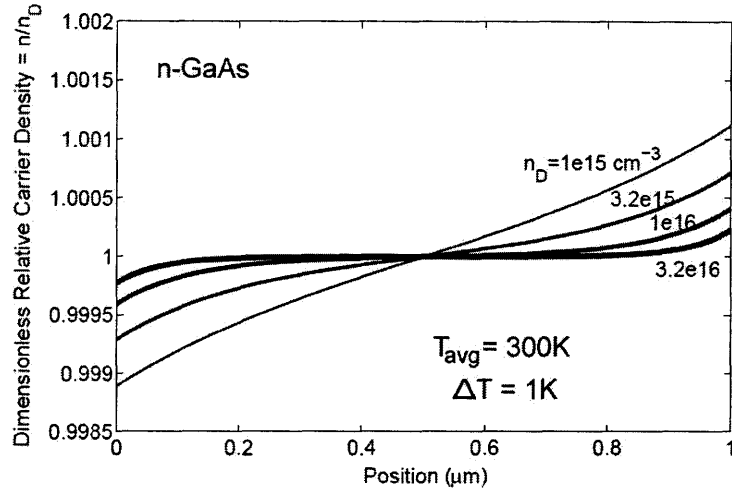


Figure 3-5: Carrier density profiles for Seebeck effect at 300K with  $\Delta T = 1\text{K}$ . As the ionized dopant concentration is increased, the greater density of free carriers serves to screen the edge effects on shorter length-scales. Just as in the determination of the Debye length, this screening length is the result of competition between drift and diffusion. It is therefore intuitive that the characteristic screening length depends on carrier density in the same way as does the Debye length ( $\propto 1/\sqrt{n_D}$ ).

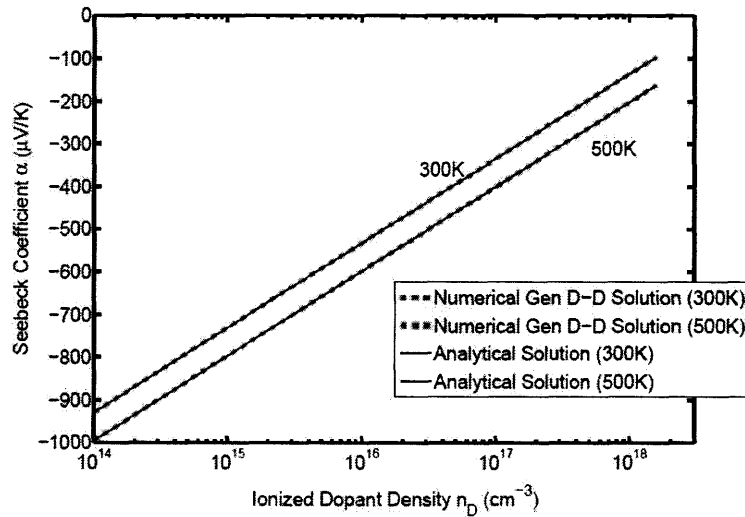


Figure 3-6: The numerical calculations for the Seebeck coefficient done via the generalized drift-diffusion framework explained in this chapter agree with the analytical results from the differential-conductivity method of Section 1.2.3 across a wide range of doping concentrations and temperatures. The calculations in the above figure are for n-GaAs, and express agreement in the Boltzmann limit. We will discuss higher ionized dopant concentrations in Section 3.5.1.

a semiconductor. All of the physical laws of electrostatics and classical transport are satisfied.

The results of these self-consistent calculations agree with the analytical results of the differential-conductivity calculation from Section 1.2.3, as shown in Figure 3-6. Recall that there was nothing incorrect in the old model, but since it gave us only the ratio between the applied  $\nabla T$  and the resulting  $\nabla E_F$  without regard to the location of the band-edge, we could not have an intuitive picture of the density profile. With the generalized drift-diffusion solution, we may readily examine the entire carrier density profile, as in Figure 3-2, and band diagram, as in Figure 3-4.

This generalized drift-diffusion framework can be readily expanded to the problem of simulating inhomogeneous thermoelectric elements, including devices with smoothly graded electronic properties as well as abrupt junctions. For example, a thermoelectric element with a graded donor concentration would be solved by the same numerical procedure, but with a different fixed-charge density entering into Poisson's equation and an average carrier density appropriately adjusted to maintain overall charge neutrality.

The additional detail provided by this model over the differential-conductivity model also yields at least one experimentally-verifiable prediction. For example, since the model gives us the profile of the conduction band-edge and the band-structure of materials shouldn't change appreciably across a semiconductor under a small temperature difference, the model also tells us the profile of the vacuum level. The model predicts that for a lightly-doped semiconductor with large  $|\alpha|$ , the majority of the difference in  $E_F$  induced by an applied temperature difference should come from the chemical potential gradient. If metal contacts are attached to ends of an appropriately-doped sample and a temperature difference  $\Delta T$  is applied between them, then a electric-potential probe sufficiently close to the surface (a distance  $d$  much less than the sample length  $L$  to avoid the fringe  $\vec{E}$ -field) should measure sharp jumps within a few Debye-lengths of the contacts, comprising all but  $86 \mu\text{V/K}$  of the total Seebeck voltage  $\Delta V = \alpha\Delta T$ . A qualitative diagram of this experimental prediction is shown in Figure 3-7.

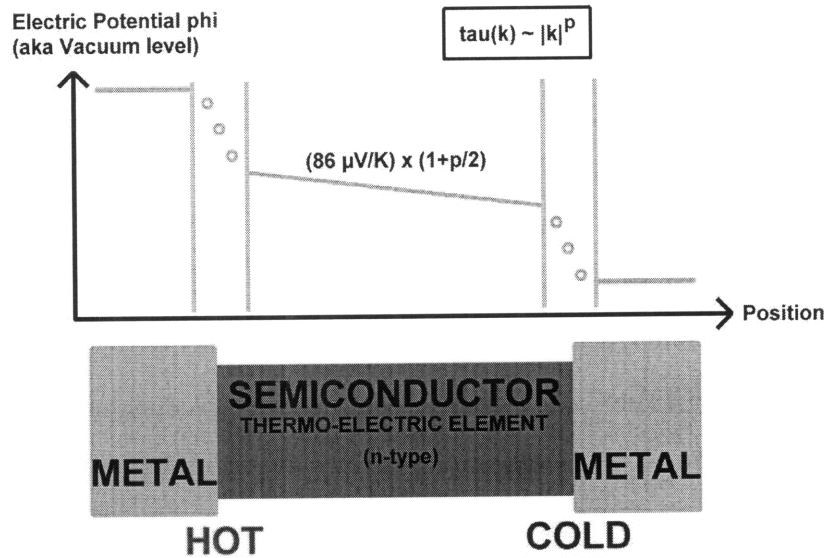


Figure 3-7: Qualitative diagram of electric potential as a function of space near the surface of a thermoelectric element under a small applied temperature gradient. All but  $86 \mu\text{V}/\text{K}$  of the potential drop should be found within a few Debye lengths of the contacts.

Finally, one additional detail may be relevant for design, though this question has not been fully explored. The deviations from quasi-neutral density at the edges of the device change the energy difference  $E_C - E_F$  at the edges of a thermoelectric element under  $\Delta T$  from the gap that would be predicted assuming a uniform density profile. Attempts to model the ends of each element to predict electrically parasitic contact resistances may need to take this non-uniformity of carriers into account. When a large number of such elements are placed electrically in series and thermally in parallel, as is done commonly in power-generating applications, the importance of contact resistance to overall system efficiency could make the details in this model relevant to design.

### 3.2.3 The Various Relaxation Times

Before extending the Seebeck solution to conditions of practical interest to the thermoelectric community (i.e. inclusion of scattering exponents, consideration of materials

doped outside the Boltzmann limit and materials with more complex bandstructures), we pause to clarify the origins of the various relaxation times  $\tau$  we have used to this point. In Chapter 2 we began by making the RTA in (2.1) with a relaxation time  $\tau$ , which at times was a constant timescale  $\tau_0$  and at other times was taken to be dependent generally on the momentum-state under consideration  $\tau(\vec{k})$ . On the other hand, in our derivation of the Momentum Balance Equation (3.9) from Section 3.2.1, a combination of somewhat opaque algebraic manipulations permitted us to define a momentum-relaxation time  $\tau_m$ , which we have thus far considered to be a constant timescale for the entire system and independent of any momentum-dependence.

While these two quantities,  $\tau$  and  $\tau_m$ , both describe the timescale for relaxing deviations of the distribution  $f(\vec{x}, \vec{k})$  to a state with no current flowing, they are in fact conceptually distinct.

To see the difference, we return to the original Boltzmann Transport Equation as expressed in Section A.1, where the right-hand side is taken to be an approximate statistical closure to an infinite hierarchy of equations which describe the classical interaction of successively larger numbers of particles simultaneously. When perturbations due to impurities or phonons dominate these interactions, the right-hand side of the BTE is well-described by scattering rates:

$$\hat{S}[f(x, k, t)] = - \sum_{k'} f(x, k, t) S(k, k') (1 - f(x, k', t)) + \sum_{k'} f(x, k', t) S(k', k) (1 - f(x, k, t)) \quad (3.18)$$

Here,  $S(k, k')$  is the probabilistic rate for a given carrier at  $k$  to be scattered to  $k'$ , assuming the state  $k$  contains a carrier and  $k'$  does not. The  $f$  and  $1 - f$  terms are present to complete the above expression for the observed rate of net in-scattering to a state  $(x, k)$  at time  $t$ . Since  $f$  describes only the distribution of carriers in phase-space,  $S$  contains all of the remaining information about the scatterers present in the system. If the distribution of carriers is dilute, the Pauli-blocking terms can be omitted, leaving:

$$\hat{S}[f(x, k, t)] = - \sum_{k'} f(x, k, t) S(k, k') + \sum_{k'} f(x, k', t) S(k', k) \quad (3.19)$$

Since by definition, the equilibrium occupancy function is stationary under scattering processes [23], by writing  $f = f_0 + \delta f$ , we may write simply:

$$\hat{S}[f(k)] = - \sum_{k'} \delta f(k) S(k, k') + \sum_{k'} \delta f(k') S(k', k) \quad (3.20)$$

The RTA, as we expressed it before, suggests that the value of the occupancy function  $f$  at any point in  $k$ -space relaxes back to its equilibrium value with some particular timescale  $\tau$ , which may ( $\tau \rightarrow \tau(\vec{k})$ ) or may not ( $\tau \rightarrow \tau_0$ ) depend explicitly on the particular location in  $k$ -space. To justify this, we need to assume something more about the detailed scattering rate  $S(\cdot, \cdot)$ , which amounts to an assumption about the interaction of the carriers with the scatterers.

The exact condition for the validity of the RTA then is simply that  $S(\cdot, \cdot)$  permits the net result of scattering to be analytically described by a relaxation time  $\tau(\vec{k})$  which, while possibly dependent on the location in  $k$ -space under consideration, must *not* be dependent on the occupancy of that state  $f(k)$  or any other state  $f(k')$ . That is, the occupancy  $f$  can be pulled out of the  $\sum_{k'}$  above:

$$\hat{S}[f(k)] = \sum_{k'} \left[ \delta f(k') S(k', k) - \delta f(k) S(k, k') \right] \rightarrow \hat{S}[f(k)] = \frac{-\delta f(k)}{\tau(\vec{k})} \quad (3.21)$$

Examples of such conditions under which the RTA as expressed are valid include when the dominant scattering process is isotropic or elastic [23]. For isotropic scattering, in which we require that the scattering rate  $S(\vec{k}, \vec{k}')$  to any state  $\vec{k}'$  is independent of the direction of  $\vec{k}'$ , so in particular  $S(k, k') = S(k, -k')$ . We notice first that by definition the equilibrium occupancy must satisfy a detailed balance of scattering rates, so that  $f_0(k)S(k, k') = f_0(k')S(k', k)$ . Then, since all of the deviations from equilibrium  $\delta f(k)$  which we've considered in Chapter 2 or Chapter 3 are odd in the

direction of deviation, we have

$$\begin{aligned}
\hat{S}[f(k)] &= \sum_{k'} \left[ \delta f(k') S(k', k) - \delta f(k) S(k, k') \right] \\
&= -\delta f(k) \sum_{k'} S(k, k') \left[ 1 - \frac{\delta f(k') f_0(k)}{\delta f(k) f_0(k')} \right] \\
&= -\delta f(k) \sum_{k'} S(k, k') \\
&= -\delta f(k) \tau^{-1}(k)
\end{aligned} \tag{3.22}$$

where the last line can be interpreted as indicating that the relaxation time  $\tau$  is just the inter-scattering time [23], which explains the recovery of the Drude values for the transport coefficients in Chapter 2 for the case of  $\tau \rightarrow \tau_0$ .

The proof that the elastic scattering condition ( $S(k, k') \rightarrow 0$  unless  $k^2 = k'^2$ ) is sufficient to justify the RTA can be found in texts on carrier transport in semiconductors [23]. The main point is that the validity of the RTA rests on the microscopic details of the dominating scattering process(es) and that the timescale for relaxation of the occupancy function toward  $f_0$  is set by the scattering rates to the other available states in  $k$ -space.

On the other hand, in each Balance Equation we considered the relaxation of a single quantity defined by a weighted average of the occupancy function. In the MBE, the weighting  $\hbar \vec{k}/m^*$  was chosen to make the quantity under consideration the rate of flux of carriers  $\vec{J}/q$ . The right-hand side of the MBE:

$$\begin{aligned}
&\int_{\mathbb{B}} d^3 \vec{k} \left( \frac{\hbar(\vec{k} \cdot \hat{e}_j)}{m^*} \left[ \hat{S}f(\vec{x}, \vec{k}, t) \right] \right) \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \int_{\mathbb{B}} d^3 \vec{k}' k_j \left[ f(k') S(k', k) - f(k) S(k, k') \right] \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} \int_{\mathbb{B}} d^3 \vec{k}' (k'_j - k_j) \left[ f(k) S(k, k') \right] \\
&= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3 \vec{k} (-k_j) f(k) \int_{\mathbb{B}} d^3 \vec{k}' S(k, k') \left( \frac{k_j - k'_j}{k_j} \right)
\end{aligned} \tag{3.23}$$

was used to define a quantity  $\tau_m$  as follows (taking the additional assumption of dilute

occupancy for simplicity):

$$\frac{1}{\tau_m(k)} \equiv \int_{\mathbb{B}} d^3 \vec{k}' \left[ S(k, k') \left( \frac{k_j - k'_j}{k_j} \right) \right] \quad (3.24)$$

Here, we have taken advantage of the fact that we do not need to describe relaxation for each  $k$ -state individually, but only for some weighted aggregate sum over all  $k$ -states, to interchange the variable names and their order of summation. In essence, instead of asking about each state independently, which requires knowledge of both in-scattering and out-scattering types of events, we simply note that each scattering event has an origin and a destination and do our accounting by summing all out-scattering events over the set of possible originating  $k$ 's. This trick, which we have no hope of playing when we want the point-wise detail of the RTA, permits us to describe momentum relaxation by the quantity  $\tau_m(k)$  independent of the distribution  $f$ . We recognize  $\tau_m(k)$  as simply the total scattering rate for a carrier at  $k$  weighted by the importance of such scattering events to relaxation of momentum.

The final form of  $\tau$  we encountered in the MBE was defined as:

$$\left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k \equiv \frac{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*} \frac{1}{\tau_m(k)}}{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*}} \quad (3.25)$$

This quantity, which we refer to as the momentum-relaxation time, is simply interpreted as the timescale over which local deviations from zero current are relaxed by scattering.

The particular forms of the deviations in the distribution function given by the RTA with  $\tau = \tau_0$  were found in Section 2.4.1, Section 2.4.2, and Section 2.4.3. For drift and  $\nabla n$ -driven diffusion currents (taken without loss of generality to be in the  $\hat{k}_j$  direction), we find that:

$$\left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k = \frac{\int_{\mathbb{B}} d^3 \vec{k} f_0(k) (1 + \epsilon k_j) \frac{\hbar k_j}{m^*} \frac{1}{\tau_m(k)}}{\int_{\mathbb{B}} d^3 \vec{k} f_0(k) (1 + \epsilon k_j) \frac{\hbar k_j}{m^*}} \quad (3.26)$$

where  $\epsilon$  is a small constant independent of  $k$ . We see that when  $\tau_m(\vec{k})$  is a function

only of the magnitude  $|\vec{k}|$ , the expression can be further simplified:

$$\left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k = \frac{\int_{\mathbb{B}} d^3\vec{k} f_0(k) (\epsilon k_j^2) (\tau_m(|\vec{k}|))^{-1}}{\int_{\mathbb{B}} d^3\vec{k} f_0(k) (\epsilon k_j^2)} \quad (3.27)$$

which is independent of the magnitude of the deviation  $\epsilon$ . We conclude that the value of  $\langle\langle \tau_m^{-1} \rangle\rangle$ , whatever it may be, indicates that currents arising from these two types of sources, electric fields and density gradients, are relaxed at the same rate. In fact, even if we permit  $\tau = \tau(|\vec{k}|)$ , we find that these two timescales remain identical.

Note that this last assumption reflects the independence of the scattering rates determined by non-carrier parts of the system, such as impurities and phonons, from the direction of distortions to the carrier distribution. A particularly relevant counter-example is the phenomenon of phonon drag, in which the anisotropic distribution of phonons on net impart momentum to the carriers, thereby enhancing the Seebeck coefficient. Nevertheless, we proceeded with the approximation for the sake of simplicity.

As we found in Section 2.4.3, the deviation of the distribution function under  $\nabla T \neq 0$  takes on a slightly different form:  $f \rightarrow (f_0)(1 + \epsilon_1 k_j + \epsilon_2 k^2 k_j)$ , where  $\epsilon_1$  and  $\epsilon_2$  are again small constants independent of  $k$ . Therefore the momentum-relaxation rate relevant to the MBE would, for currents driven by nonzero  $\nabla T$ , be:

$$\left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k = \frac{\int_{\mathbb{B}} d^3\vec{k} f_0(k) (\epsilon_1 k_j^2 + \epsilon_2 k^2 k_j^2) (\tau_m(|\vec{k}|))^{-1}}{\int_{\mathbb{B}} d^3\vec{k} f_0(k) (\epsilon_1 k_j^2 + \epsilon_2 k^2 k_j^2)} \quad (3.28)$$

which is in general not the same as the timescale for current decay from the other sources. In the case with no scattering exponent ( $\tau_m(|\vec{k}|) = \tau_m$ ), however, we conclude that the timescale for current decay from all three sources is equivalent. This is the detailed justification for the analysis presented in Section 3.2. In the next section, Section 3.3, we consider systems with nonzero scattering exponents; the simplest solution calls for a mixture of the MBE and RTA results, and our work in both Chapter 2 Chapter 3 will be rewarded.

In summary, several distinct relaxation timescales exist in the Relaxation-Time



Approximation and Momentum Balance Equation approaches to solving the BTE. By returning to the basic expression describing relaxation in terms of the point-to-point scattering rate  $S(k, k')$ , we see that not only do  $\tau$  and  $\tau_m$  represent distinct physical quantities, their usage alone is justified by two different sets of assumptions about the relevant underlying scattering processes. Since we can solve for the Seebeck coefficient by the MBE approach, as we have in Section 3.2, we are not forced into making the RTA to calculate the Seebeck coefficient microscopically, regardless of its validity for a particular system.

### 3.3 Seebeck with Energy-Dependent Relaxation Time

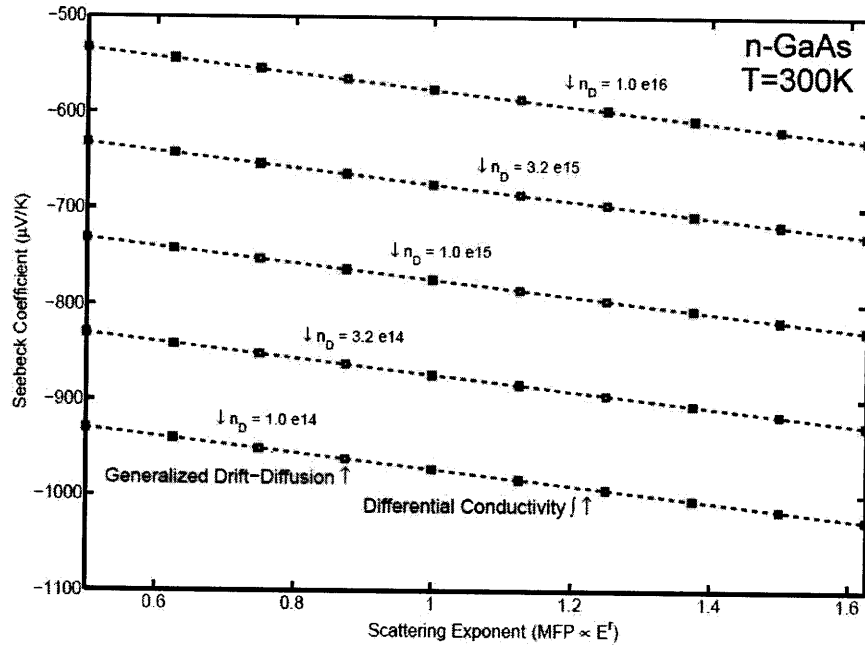


Figure 3-8: Plot showing agreement across a range of doping densities between generalized drift-diffusion model and differential conductivity model for Seebeck coefficient as a function of scattering parameter  $r$ . The scattering parameter is defined as the exponent of energy-dependence in mean-free-path:  $\lambda(E) = \lambda_{th} \left( \frac{E}{k_B T} \right)^r$ . All calculations are for n-GaAs at 300K.

The Seebeck effect is stronger in semiconductors for which the dynamics of carrier distribution relaxation leads to slower return to equilibrium for higher-energy carriers.

To see why, let us re-examine Figure 2-3. As we saw in (2.15), the diffusion current driven by temperature gradients takes place on average at larger  $|\vec{k}|$  than currents driven by electric fields or density gradients. In fact, for small wave-vectors, the contribution to diffusion current points toward higher temperatures. Simply put, if we permit slower relaxation of higher-energy carriers, they will become more important for the net transport of charge, so that the Soret coefficient will become relatively larger than the Diffusion coefficient or conductivity. Then, as we saw in Section 3.2, the amount of current that will need to be offset by the combination of  $\nabla n$ -driven diffusion current and drift current will be larger, causing more carriers to pile up on the cold side and more depletion to take place on the hot-side. The net result is a larger Seebeck coefficient.

With this intuitive explanation in mind, let us quantify this improvement. Our strategy is straightforward, and since we make use of results from Chapter 2, its validity rests on assumptions underlying the RTA. We begin with the steady-state Momentum Balance Equation, keeping the scattering terms in their general form:

$$\begin{aligned} \frac{nk_B}{m^*}(-\nabla T) + \frac{k_B T}{m^*}(-\nabla n) + \frac{qn}{m^*}E_z &= \frac{\vec{J}}{q} \left\langle \left\langle \frac{1}{\tau_m} \right\rangle \right\rangle_k \\ \frac{nk_B}{m^*}(-\nabla T) + \frac{k_B T}{m^*}(-\nabla n) + \frac{qn}{m^*}E_z &= \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} \left[ k_z \frac{1}{\tau_m(k)} f(k) \right] \end{aligned} \quad (3.29)$$

where we have taken all gradients to be in the  $\hat{z}$  direction and  $f$  here refers to the phase-space density.

Since we used the RTA in Chapter 2 to estimate the deviations in the distribution function  $\delta f$  due to each of the three sources present, we can write  $f = f_0 + \delta f_{\vec{E}} + \delta f_{\nabla n} + \delta f_{\nabla T}$  and substitute this expression into the integral for  $\langle\langle 1/\tau_m \rangle\rangle$ :

$$\frac{nk_B}{m^*}(-\nabla T) + \frac{k_B T}{m^*}(-\nabla n) + \frac{qn}{m^*}E_z = \frac{\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} \left[ k_z \frac{1}{\tau_m(k)} (f_0(k) + \delta f_{\vec{E}} + \delta f_{\nabla n} + \delta f_{\nabla T}) \right] \quad (3.30)$$

Note that for isotropic momentum-relaxation time  $\tau_m(|\vec{k}|)$ , which we showed in Section 3.2.3 is much less stringent than requiring isotropic point-to-point scattering rates  $S(k, k')$ , the contribution of  $f_0$  is zero because the integrand is odd in  $k_z$ . Now

we observe that each of these deviation terms is linear in exactly one of the driving gradient quantities. Grouping terms according to sources:

$$\begin{aligned}
& \left[ \frac{nk_B}{m^*}(-\nabla T) + \frac{-\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} k_z \frac{1}{\tau_m(k)} \delta f_{\nabla T} \right] \\
& + \left[ \frac{k_B T}{m^*}(-\nabla n) + \frac{-\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} k_z \frac{1}{\tau_m(k)} \delta f_{\nabla n} \right] \\
& + \left[ \frac{qn}{m^*} E_z + \frac{-\hbar}{m^*} \int_{\mathbb{B}} d^3\vec{k} k_z \frac{1}{\tau_m(k)} \delta f_{\vec{E}} \right] = 0
\end{aligned} \tag{3.31}$$

We are left with an equation that strongly resembles the Momentum Balance Equation, but with transport coefficients which depend on  $\tau_m(|\vec{k}|)$ . Since each of the terms in brackets has units of current density, we may regard this equation as balancing three distinct currents, each proportional to an independent source gradient. Recalling that in Chapter 2 we calculated the current in the presence of power-law relaxation times ( $\tau \propto |\vec{k}|^p$ ) and used them to calculate new transport coefficients ( $\sigma(p), D(p), S(p)$ ). Applying these results and cancelling the ratio of Gamma functions that is common to all three, we have:

$$\frac{\Gamma\left(\frac{5+p}{2}\right)}{\Gamma\left(\frac{5}{2}\right)} \left[ \left(1 + \frac{p}{2}\right) \frac{nk_B}{m^*}(-\nabla T) + \frac{k_B T}{m^*}(-\nabla n) + \frac{qn}{m^*} E_z \right] = 0 \tag{3.32}$$

The net result, we see, is that the Soret coefficient is multiplied by  $(1 + \frac{p}{2})$ , which in the Seebeck effect requires that the electric field and density gradient combine to produce  $(1 + \frac{p}{2})$  times as much opposing current density at each point. The resulting source-specified current-density profile is shown for  $p = 5$  (impurity scattering [24]) in Figure 3-9 and for  $p = -1$  (acoustic phonon scattering [24]) in Figure 3-10. The situation qualitatively resembles that from Figure 3-3 and so the intuitive picture of how the drift and  $\nabla n$ -diffusion currents combine self-consistently related in Section 3.2 need not be repeated.

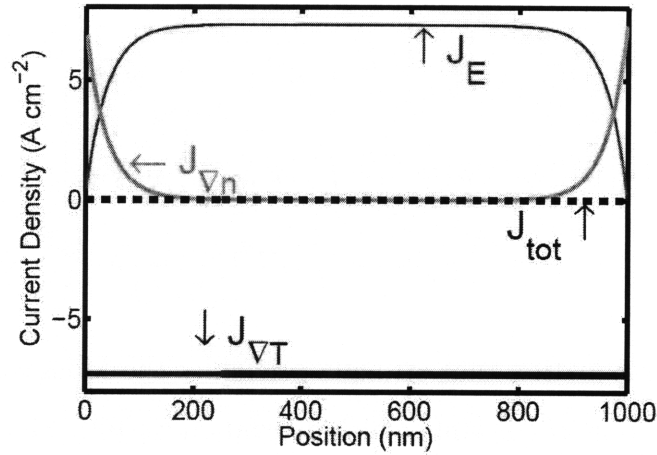


Figure 3-9: The source-specified current density profile for the Seebeck effect in a material with scattering time  $\tau \propto |\vec{k}|^5$ , indicative of ionized-impurity scattering [24]. All other parameters are identical to Figure 3-3, including  $\tau_{th} = \tau(|\vec{k}| = \sqrt{m^*k_B T}/\hbar)$ .

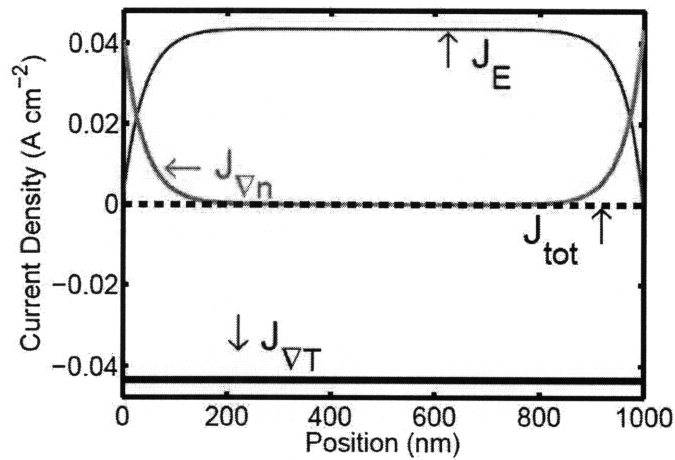


Figure 3-10: The source-specified current density profile for the Seebeck effect in a material with scattering time  $\tau \propto |\vec{k}|^{-1}$ , indicative of ionized-impurity scattering [24]. All other parameters are identical to Figure 3-3, including  $\tau_{th} = \tau(|\vec{k}| = \sqrt{m^*k_B T}/\hbar)$ .

### 3.4 Decomposing the Seebeck Effect

Another benefit of the model is that we can separate the contributions to the Seebeck coefficient. In the Seebeck effect, an applied temperature difference  $\Delta T$  gives rise to a difference in the electro-chemical potential  $\Delta E_F = \Delta\mu + q\Delta\phi$  across the device, which can be decomposed into the following three contributions:

- The difference in the chemical potential due to the difference in density across the device:  $(\Delta E_F)_n = \left(\frac{\partial\mu}{\partial n}\right)_T \Delta n$
- The difference in the chemical potential due to the difference in temperature across the device:  $(\Delta E_F)_T = \left(\frac{\partial\mu}{\partial T}\right)_n \Delta T$
- The difference in the electric potential energy, given in terms of the line-integral of the electric field across the device:  $(\Delta E_F)_\phi = \int_{x_{hot}}^{x_{cold}} dx' \left[-q\vec{E}(x')\right]$

Note that the first two contributions sum to give the entire change in the chemical potential because the Seebeck effect is defined in the limit of infinitesimal differences in all quantities across a given device, including  $\Delta\mu$ .

While this conceptual decomposition does not require a microscopic derivation of the sort found in Section 3.2, in order to quantify the contributions, we must have explicit knowledge of the electric field profile not provided by the derivation in Section 1.2.3.

The results from Section 3.2 permit us to apply this decomposition to the analytical expression for the Seebeck coefficient in semiconductors with carrier concentrations in the Boltzmann limit. Taking  $p$  to be the power-law dependence of the relaxation time on momentum, we may translate (1.33) as follows, using  $r = p/2$  for parabolic bandstructures:

$$\alpha = \frac{k_B}{q} \left[ \frac{5}{2} + \frac{p}{2} + \log\left(\frac{N_C}{n}\right) \right] \quad (3.33)$$

Since the electric potential can be defined up to an additive constant as  $q\phi = E_C$ , and our interest is only in differences in  $q\phi$  and  $\mu$  anyway, we may identify the chemical potential as  $\mu = E_F - q\phi = E_F - E_C$ . In the dilute limit of an n-type semiconductor,

a simple expression for it can be derived from the common expression for the density:

$$n = N_C \exp \left[ \frac{-(E_C - E_F)}{k_B T} \right] \text{ where } N_C \equiv 2 \left( \frac{m^* k_B T}{2\pi \hbar^2} \right)^{3/2} \quad (3.34)$$

$$\implies \mu(x) = k_B T(x) \log(n(x)/N_C(x))$$

From here, we can calculate the two contributions to the difference in chemical potential:

$$\begin{aligned} \left( \frac{\partial \mu}{\partial n} \right)_T &= \frac{k_B T}{n} \\ \implies (\Delta E_F)_n &= k_B T \frac{\Delta n}{n} \\ \left( \frac{\partial \mu}{\partial T} \right)_n &= k_B \log(n/N_C) - k_B T \frac{\partial}{\partial T} \log(N_C/n) \\ &= k_B \log(n/N_C) - \frac{k_B T}{N_C} \frac{\partial N_C}{\partial T} \\ &= k_B \log(n/N_C) - \frac{k_B T}{N_C} \frac{N_C}{T} \frac{3}{2} \\ &= k_B \left( \log(n/N_C) - \frac{3}{2} \right) \\ \implies (\Delta E_F)_T &= k_B \left( \log(n/N_C) - \frac{3}{2} \right) (\Delta T) \end{aligned} \quad (3.35)$$

We pause to note that ignoring the dependence of  $N_C$  on  $T$ , or equivalently the explicit dependence of  $\mu$  on  $T$  has been a source of sufficient confusion to the community to warrant literature dedicated to its resolution [25, 14]. Moving along, since  $\alpha = \frac{\Delta E_F/|q|}{\Delta T}$ , we know that the first contribution  $\alpha_T$  is (for the electrons in an n-type semiconductor):

$$\alpha_T = \frac{k_B}{q} \left[ \frac{3}{2} + \log(N_C/n) \right] \quad (3.36)$$

leaving

$$\alpha_n + \alpha_\phi = \frac{k_B}{q} \left[ 1 + \frac{p}{2} \right] \quad (3.37)$$

To further decompose the share of Seebeck into the  $n$  and  $\phi$  contributions, we must examine the real-space solution from Section 3.2. Since almost all of the density difference between the two ends of the device can be attributed to the regions within a Debye length ( $\lambda_D$ ) of the edges, as seen in Figure 3-2. Meanwhile, in the center

of a device with length  $L \gg \lambda_D$ , to a close approximation  $\nabla n \rightarrow 0$ . There, the Momentum Balance Equation reads:

$$\frac{\Gamma\left(\frac{5+p}{2}\right)}{\Gamma\left(\frac{5}{2}\right)} \left(1 + \frac{p}{2}\right) \frac{nk_B}{q} \mu (-\nabla T) + \frac{\Gamma\left(\frac{5+p}{2}\right)}{\Gamma\left(\frac{5}{2}\right)} n\mu \vec{E} = 0 \quad (3.38)$$

and the electric field can be directly solved for, giving us  $\alpha_\phi$ :

$$\begin{aligned} \frac{|\vec{E}|}{|\nabla T|} &= \left(1 + \frac{p}{2}\right) \frac{k_B}{q} \approx \left(1 + \frac{p}{2}\right) \times 86\mu\text{V/K} \\ \Rightarrow q\Delta\phi &= \int_{x_{hot}}^{x_{cold}} dx' \left[-q\vec{E}(x')\right] \\ &= \left(1 + \frac{p}{2}\right) (-k_B)(\Delta T) \\ \Rightarrow \alpha_\phi &= \frac{k_B}{q} \left(1 + \frac{p}{2}\right) \end{aligned} \quad (3.39)$$

which accounts for the entire remainder of the expression  $\alpha_n + \alpha_\phi$ . Note that our common observation of n-type semiconductors with  $\alpha > O(86\mu\text{V/K})$  implies that this contribution alone cannot account for the entire Seebeck effect, as the field-only picture in Figure 3-1 may initially suggest. The origin of the  $p$ -dependence in Figure 3-7 should also now be clear.

From here we can see the full decomposition of the Seebeck effect in an n-type semiconductor in the Boltzmann limit. The first contribution,  $\alpha_T$ , is large and is the part which becomes larger as the carrier concentration is reduced. The other two contributions make up the remainder, including the impact of the scattering parameter, but trade off their relative importance in various parts of the device. Since the contribution of  $\alpha_n$  is only significant within  $\lambda_D$  of the end of the device, the exact decomposition of the Seebeck coefficient in any particular sample is dependent on the size of the sample. When  $L \gg \lambda_D$ ,  $\alpha_\phi$  is the dominant remaining contribution, but when  $L \ll \lambda_D$  in open-circuit,  $\alpha_n$  dominates. Still, for sufficiently low doping concentrations,  $\alpha_T$  is much more important overall.

## 3.5 More Realistic Material Systems

To this point, we have made a number of simplifications to permit analytical and computational solutions. Some of these simplifications are inconsistent with common material systems found in viable thermoelectric generators and coolers.

### 3.5.1 Optimal Doping for Power-Generation

While most of the simple results from Chapter 2 and Section 3.2 were calculated in the Boltzmann limit, where the carriers form a dilute gas and the quantum statistics of the carriers are irrelevant, the figure-of-merit  $ZT$  is typically optimized at doping concentrations which result in an equilibrium Fermi level within a few  $k_B T$  of the band-edge.

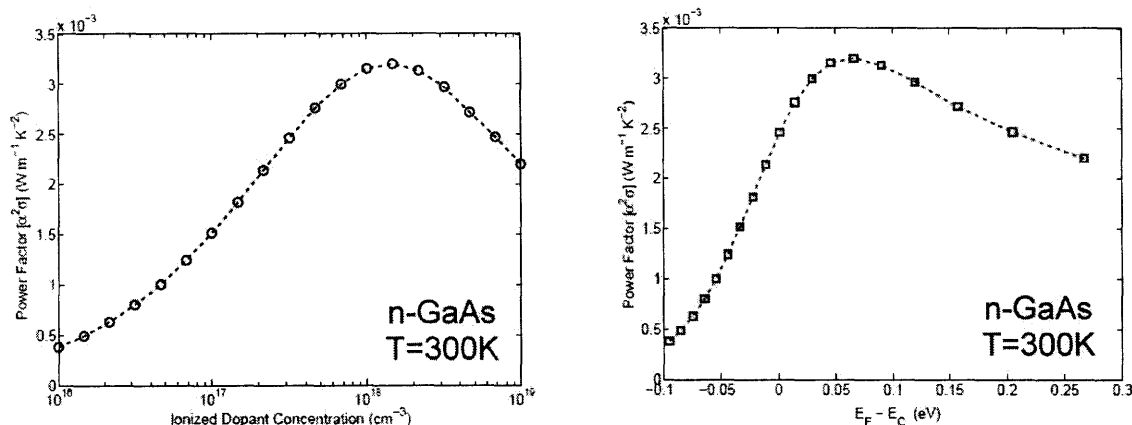


Figure 3-11: The so-called thermoelectric power factor ( $\alpha^2 \sigma$ ) which enters directly into  $ZT$  as seen in (1.4), experiences a maximum as a function of doping near the concentration required for the equilibrium Fermi level to fall near the band-edge. These plots express the results of the computational generalized drift-diffusion solver, using results for the transport coefficients from Section 2.5. Left: Power factor versus ionized dopant density. Right: Power factor versus equilibrium  $E_F(n_D)$  measured relative to  $E_C$ .

Although this situation is obviously incommensurate with the Boltzmann approximation, our general strategy remains valid, since the validity of the Boltzmann Transport Equation and Poisson's equation remain intact. For the same reason that we saw when examining  $k$ -dependent relaxation times in Section 3.3, we may use the same



framework but substituting the proper values of  $D$ ,  $S$ , and  $\sigma/q$  from Section 2.5 into the generalized drift-diffusion equation (3.14). As before, we are justified in doing so because with or without the Boltzmann limit, when  $\tau_m(\vec{k}) = \tau_m(|\vec{k}|)$ , the scattering term in the MBE can be broken apart into currents dependent on the deviations and rearranged to give an equation expressing a balance of fluxes based on the present gradients ( $\nabla n$ ,  $\nabla T$ , and  $\vec{E}$ ).

In the Boltzmann limit, there are relatively few carriers present for conduction. The power factor  $\alpha^2\sigma$  increases with carrier concentration because  $\alpha$  only depends on the log of the carrier density while the  $\sigma$  depends on  $n$  linearly.

The transport coefficients deviate from their asymptotic dilute values as shown in Section 2.5. Since the ratio of the Soret coefficient to the diffusivity and conductivity decays to zero as doping is increased and the equilibrium  $E_F$  is pushed into the band, the strength of the gradients  $\nabla n$  and  $\vec{E}$  required to drive  $\nabla n$ -diffusion and drift respectively decay to zero as well. In the language of Section 3.4, this means that  $\alpha_n$  and  $\alpha_\phi$  decay as doping becomes more degenerate. Meanwhile, as the density-of-states ( $D(E) \propto E^{1/2}$ ) near  $E_F$  becomes flatter, the change in the chemical potential  $E_F - E_C$  required to maintain a fixed carrier concentration in the presence of a temperature difference across the sample also disappears, and  $\alpha_T$  with it. Hence, we expect to find a doping concentration between the Boltzmann and degenerate limits which maximizes the power factor  $\alpha^2\sigma$ , as shown in Figure 3-11.

### 3.5.2 A Note on Complex Bandstructures

The high- $ZT$  materials most commonly found in thermoelectric generators and coolers (see Figure 3-12) typically do not have the simple spherical, parabolic bandstructures we have assumed until now. In fact, their band-edges are typically composed of numerous nearly-degenerate valleys. For example, the commercially-popular material  $\text{Bi}_2\text{Te}_3$  has a 6-fold valley-degeneracy in both the conduction and valence band-edges [4]. To see why, consider that in the Boltzmann limit, a larger band-edge density-of-states implies that when a temperature difference is applied, the chemical potential must change by a larger amount to maintain a fixed density, so  $\alpha_T$  is larger. This

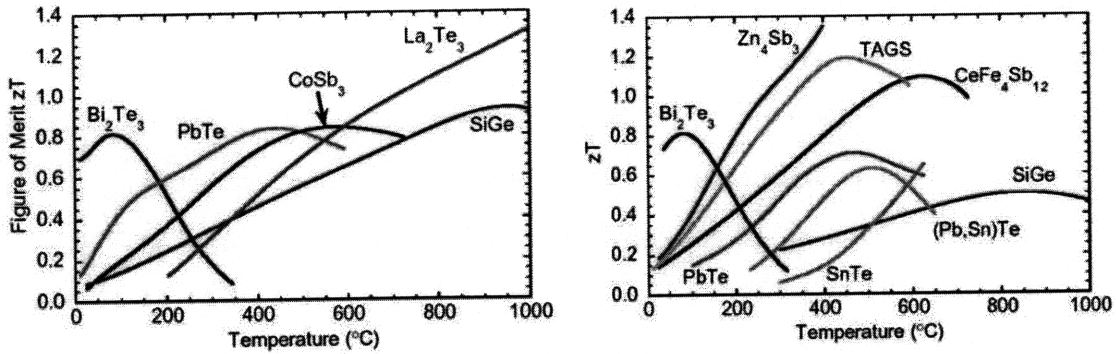


Figure 3-12: Material ZTs with temperature for various technologically-relevant materials. Left: n-type. Right: p-type. Image taken from [20].

intuition is consistent with the recent theoretical result from Humphrey and Linke [9] that the density-of-states which optimizes  $\alpha^2\sigma$  would be a delta-function at the band-edge:  $D(E) = \delta(E - E_C)$ .

In spite of the technological relevance, we have not yet implemented a model for materials whose band-edges consist of multiple valleys. To properly simulate such materials, we must incorporate an additional timescale into the problem, determined by the inter-valley scattering rate. Returning to (2.3), we see that the initial result from the Relaxation Time Approximation suggests that the solution to the BTE under density gradient, for example, for carriers in valleys located far from the zone-center  $\Gamma$  will have their occupancy determined by the equilibrium occupancy from a distance  $\frac{\hbar\tau}{m^*} \vec{k}_{C-\Gamma}$  where  $C$  is the location of the band minimum. The blind application of this result would suggest that band minima far from  $\Gamma$  would have huge diffusion coefficients in the direction of  $\hat{k}_{C-\Gamma}$  which were determined primarily by  $|\vec{k}_{C-\Gamma}|$  and not the relevant dynamical effective mass.

Alternatively, we could take each valley to be centered at  $\Gamma$  for the purposes of calculating a transport coefficient ( $S$ ,  $D$ , or  $\sigma$ ), then try to combine these results. Consider Si, in which the six equivalent valleys have the same longitudinal and transverse effective masses, but for any given direction have distinct dynamical effective masses. The Boltzmann-limit calculations performed assuming  $C \rightarrow \Gamma$  would yield the same results as before ( $S = \frac{nk_B}{q} \frac{q\tau}{m^*}$ ,  $D = \frac{k_B T}{q} \frac{q\tau}{m^*}$ , and  $\sigma = n \frac{q\tau}{m^*}$ ) with  $m^*$  referring

to the dynamical effective mass. If we proceed with calculating a Seebeck coefficient for each valley, when we looked to combine these results we would find that they suggest occupancies in distinct valleys at the same location in real-space reflective of different values of the local  $E_F$ . These differences in  $E_F$  between valleys would have been permitted to persist only because we did not consider inter-valley scattering.

Resolving these issues to create an acceptable model of thermoelectric transport in multi-valley semiconductors is clearly an important priority going forward in the development and application of the generalized drift-diffusion technique.

### 3.6 The Microscopic Peltier Effect

Until now, we have developed tools to describe semiconductors in the presence of thermal gradients under electrical open-circuit boundary conditions because this particular situation defines the Seebeck effect and may be of value in the design of thermo-electric generators. Nevertheless, thermoelectric phenomena often find application in situations more closely-resembling the setup of the Peltier effect), such as in thermoelectric coolers internally-cooled micro-electronic devices.

Recall from Figure 1-2 that the difference in Peltier coefficient across an isothermal interface is defined as the ratio of the reversible heat produced at the interface to the current flowing across the junction.

In this section, we use the balance equations to present a microscopic picture of the Peltier effect. We first derive the Energy Balance Equation and identify the Peltier heat term, noting its consistency with the notion of the difference in average transport energy. We then proceed to simulate the Peltier effect at a semiconductor homojunction using the Momentum Balance Equation. We leave discussion of heterojunctions to future work (see Section 5.3).

### 3.6.1 The Energy Balance Equation and Peltier Heat

The Energy Balance Equation follows the MBE in the balance equation heirarchy, and expresses the equality of the  $k^2$ -moment of each term in the BTE:

$$\int_{\mathbb{B}} d^3\vec{k} \frac{\hbar^2|\vec{k}|^2}{2m^*} \left[ \left( \frac{\partial}{\partial t} + \frac{\hbar}{m^*} \vec{k} \cdot \nabla + \frac{q}{\hbar} \vec{E} \cdot \nabla_k \right) f(\vec{x}, \vec{k}, t) \right] = \int_{\mathbb{B}} d^3\vec{k} \frac{\hbar^2|\vec{k}|^2}{2m^*} \left[ \hat{S}f(\vec{x}, \vec{k}, t) \right] \quad (3.40)$$

Examining each term independently:

$$\int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2|\vec{k}|^2}{2m^*} \frac{\partial}{\partial t} f \right) \equiv \frac{\partial}{\partial t} W = \frac{\partial}{\partial t} \text{tr}(\overline{W}) \quad (3.41)$$

$$\int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2|\vec{k}|^2}{2m^*} \right) \left[ \frac{\hbar\vec{k}}{m^*} \cdot \nabla f \right] = \frac{\partial}{\partial x_i} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2|\vec{k}|^2}{2m^*} \right) \left[ \frac{\hbar k_i}{m^*} f \right] \equiv \nabla \cdot \vec{F}_W \quad (3.42)$$

$$\begin{aligned} \int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2|\vec{k}|^2}{2m^*} \left[ \frac{q\vec{E}}{\hbar} \cdot \nabla_k f \right] \right) &= \frac{-q\hbar}{2m^*} \int_{\mathbb{B}} d^3\vec{k} f \frac{\partial}{\partial k_i} (|\vec{k}|^2 E_i) \\ &= \frac{-q\hbar}{2m^*} \int_{\mathbb{B}} d^3\vec{k} f (2k_i E_i) \\ &= - \int_{\mathbb{B}} d^3\vec{k} f \left( q \frac{\hbar\vec{k}}{m^*} \cdot \vec{E} \right) \\ &= -\vec{E} \cdot \vec{J} \end{aligned} \quad (3.43)$$

$$\begin{aligned} &\int_{\mathbb{B}} d^3\vec{k} \left( \frac{\hbar^2|\vec{k}|^2}{2m^*} \left[ \hat{S}f \right] \right) \\ &= \frac{\hbar^2}{2m^*} \int_{\mathbb{B}} d^3\vec{k} \int_{\mathbb{B}} d^3\vec{k}' (-|\vec{k}|^2) \left[ f(k)S(k, k')(1 - f(k')) \left( 1 - \frac{|\vec{k}'|^2}{|\vec{k}|^2} \right) \right] \\ &= \frac{-\hbar^2}{2m^*} \int_{\mathbb{B}} d^3\vec{k} \left[ f(k) |\vec{k}|^2 \frac{1}{\tau_E(k)} \right] \\ &= -(W - W_0) \left\langle \left\langle \frac{1}{\tau_E} \right\rangle \right\rangle_{k^2} \end{aligned} \quad (3.44)$$

where we have defined:

$$\frac{1}{\tau_E(k)} \equiv \int_{\mathbb{B}} d^3 \vec{k}' \left[ S(k, k') (1 - f(k')) \left( 1 - \frac{|\vec{k}'|^2}{|\vec{k}|^2} \right) \right]$$

and

$$\left\langle \left\langle \frac{1}{\tau_E} \right\rangle \right\rangle_{k^2} \equiv \frac{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*} \frac{1}{\tau_E(k)}}{\int_{\mathbb{B}} d^3 \vec{k} f(k) \frac{\hbar k_j}{m^*}} \quad (3.45)$$

See Section 3.2.3 for further clarification on  $\langle \langle \tau_E^{-1} \rangle \rangle_{k^2}$ . The simplification of the scattering term found above was taken from [23].

Combining these results, we arrive at the Energy Balance Equation:

$$\frac{\partial W}{\partial t} + \nabla \cdot \vec{F}_W - \vec{E} \cdot \vec{J} = \frac{-(W - W_0)}{\left( \left\langle \left\langle \frac{1}{\tau_E} \right\rangle \right\rangle_{k^2} \right)^{-1}} \quad (3.46)$$

Since the Peltier effect is defined under isothermal conditions, the energy density everywhere can be set to the equilibrium value  $W_0$  to which scattering drives  $W$ , and  $\frac{\partial W}{\partial t} \rightarrow 0$  as well. We are left with simply:

$$\nabla \cdot \vec{F}_W = \vec{E} \cdot \vec{J} \quad (3.47)$$

which states that the net outflow of energy density at any point is equal to  $\vec{E} \cdot \vec{J}$ . As with any local quantity, continuity of energy density requires that accumulation plus net outflow equal generation:

$$\frac{\partial W}{\partial t} + \nabla \cdot \vec{F}_W = \dot{W} \quad (3.48)$$

so we may identify  $\vec{E} \cdot \vec{J}$  with the local generation of energy density. Of course this statement is not in violation of the isothermal condition— it only requires that the locally-generated energy density go someplace outside the described system immediately to maintain the boundary condition. For the system of electronic carriers described here, the destination of this energy density is some other subsystem of the

sample, typically the phonons in the lattice though other species of excitations are possible as well.

We note that in a semiconducting resistor, where  $\vec{E} \cdot \vec{J}$  is homogeneous and positive throughout the sample, the heat which is generated is not reversible Peltier heat, but irreversible Joule heat. To separate the Peltier heat from the Joule heat, we simply subtract the contribution from the flow of electrons down an electro-chemical potential gradient ( $(\vec{J}/q) \cdot -\nabla E_F$ ), since this represents the power deposited in a resistive voltage drop.

A second way to see the need to remove this term is to recall that, as we found in (B.1), the rate of entropy generation (as opposed to entropy accumulation) is proportional to  $\vec{J} \cdot (-\nabla E_F/q)$  as well. Entropy production is irreversible, so we need to remove this component. The Peltier effect describes the reversible process of transporting of entropically-bound energy (also known as heat) from one location to another, not the conversion of free energy into heat.

One nice way to interpret these terms comes from thermodynamics, where any change in energy  $dU$  of a gas of particles can be ascribed to either heat  $TdS$  or Helmholtz free energy  $dF = dU - TdS$ . Since  $dU = TdS + PdV + E_F dN$ ,  $E_F = (\partial F/\partial N)_V$ , and we can identify the electro-chemical potential  $E_F$  at any point with the quantity of free-energy that would be added to the system by the addition of one charge carrier at that location. We recall that when a group of electrons flow down a part of the device where the band-edge is flat and  $\vec{J} \cdot \vec{E}$  is zero, the energy associated with that group is unchanged by (3.46). However, the quantity of free-energy associated with that group is reduced, indicating that some of the free energy has been entropically bound into heat, meaning entropy has been produced. On the other hand, if the group of electrons travel across a region with flat  $E_F$  but sloped band-edge ( $\vec{E} \neq 0$ ), then we have an increase in energy associated with those particles from the energy balance equation, but the free-energy associated with them stays the same. Therefore the additional energy that has entered the system must've been accompanied by sufficient entropy to call all of it heat. This heat, did not involve converting free-energy into entropically-heat and therefore irreversibly generating entropy, is the

reversible Peltier heat.

Writing the electric field as  $\vec{E} = -\nabla\phi = -\nabla E_C/q$ , we may identify the rate at which Peltier heat density enters the electronic system as:

$$Q_{\text{Peltier}} = \vec{J} \cdot \left( \frac{-\nabla E_C}{q} - \frac{-\nabla E_F}{q} \right)$$

or

$$Q_{\text{Peltier}} = \vec{J} \cdot \left( \frac{\nabla\mu}{q} \right)$$
(3.49)

As a final check on this expression for the Peltier heat density, we note that if we integrate the Peltier heat density over the space around the junction, since the sample has constant current flowing, we find:

$$\int dx' Q_{\text{Peltier}}(x') = \int dx' \vec{J}(x') \cdot \left( \frac{\nabla\mu(x')}{q} \right) = J (-\Delta(E_C - E_F)/q) \quad (3.50)$$

exactly recovering our intuition from Section 1.1 that the absolute Peltier coefficient of a material reflects the energy transported per unit charge.

It is also interesting to note that the same equations which describe the Peltier effect, when applied to a material with continuously-varying zero-current chemical potential due to a constant temperature gradient, are readily interpreted as the Thomson effect. Since in semiconductor homojunctions, all of our differences in equilibrium chemical potential are spread out by distances  $O(\lambda_D)$ , we may simply incorporate both the Peltier and Thomson effects together by considering only gradients in the chemical potential at zero current, without regard to whether its origin is a doping concentration or other local electric potential perturbation (Peltier effect) or temperature gradient (Thomson effect).

### 3.6.2 Homojunctions

Implementing the strategy for simulating the Peltier effect in  $n^+ - n$  homojunctions proceeds as outlined in Section 3.6.1. We simply solve the Momentum Balance Equation and Poisson's equation self-consistently in the presence of drift and  $\nabla n$ -diffusion

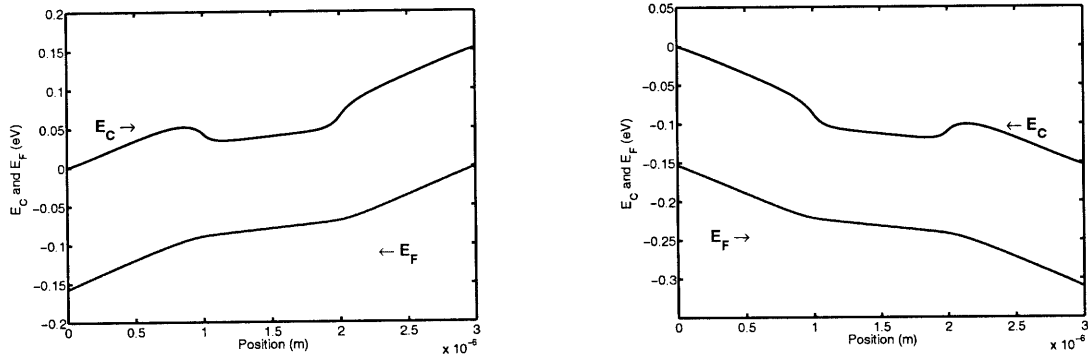


Figure 3-13: Band-edge profiles for  $n - n^+ - n$  structure at 300K with currents of  $+1 \text{ kA/cm}^2$  (left) and  $-1 \text{ kA/cm}^2$  (right) flowing. The doping concentrations of both the  $n$  ( $n_D = 1e15 \text{ cm}^{-3}$ ) and  $n^+$  ( $n_D = 4e15 \text{ cm}^{-3}$ ) remain in the Boltzmann limit. Other material parameters are taken from GaAs:  $m^* = 0.063m_e$ ,  $\epsilon_r = 10.6$ ,  $\tau_m = 0.3\text{ps}$ . The isothermal boundary condition of the Peltier effect is enforced, which is equivalent to  $\kappa \rightarrow \infty$ .

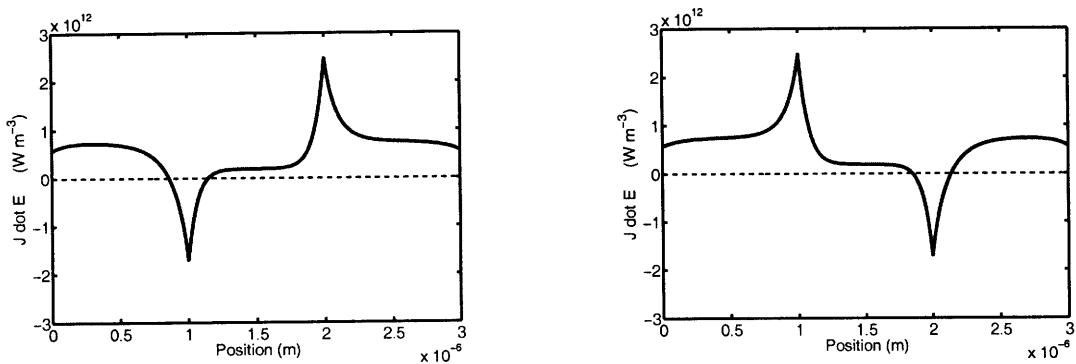


Figure 3-14: Rate of total heat density production, including both Peltier and Joule contributions, in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), some of the heat density production flips sign but some does not.



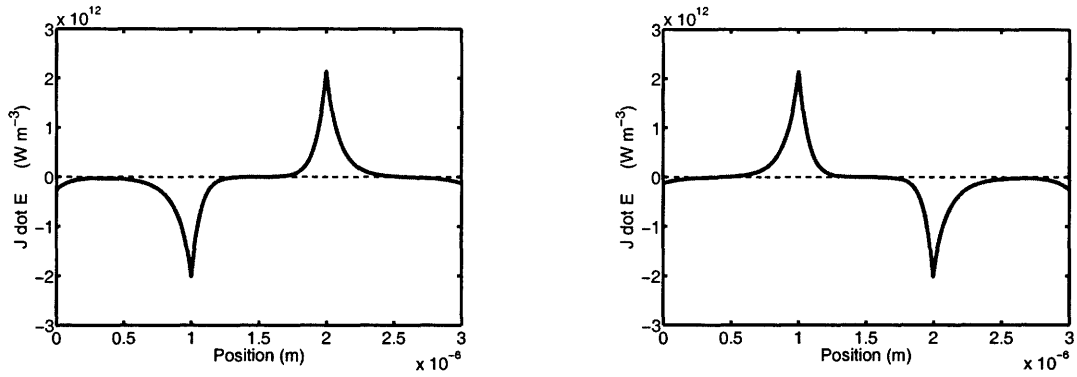


Figure 3-15: Rate of Peltier-only heat density production in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), all of the Peltier heat density production flips sign.

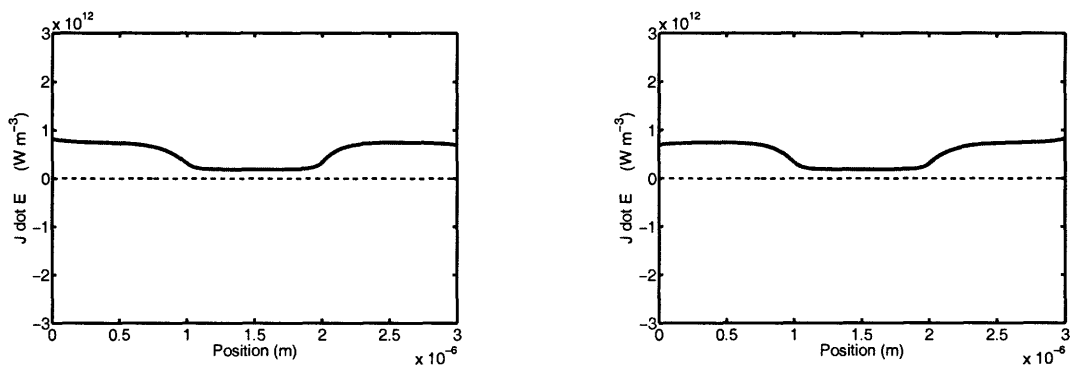


Figure 3-16: Rate of Joule-only heat density production in the structure shown in Figure 3-13. Note that as the positive current (left) is reversed (right), none of the Joule heat density production flips sign.

currents, omitting  $\nabla T$ -diffusion because of the definitional isothermal boundary condition.

We present the results of passing a current through an  $n - n^+ - n$  structure with two complementary homojunctions. The self-consistent solution that results includes the local electric field profile, which can be multiplied by the constant current being driven to find the heat production  $\vec{J} \cdot \vec{E}$ , shown in Figure 3-14. The density profile can also be used to find the chemical potential  $\mu = -(E_C - E_F)$ , whose gradient multiplies the current to give the Peltier heating shown in Figure 3-15. The chemical potential profile can then be combined with the electric field profile to calculate the electro-chemical potential  $E_F$ , whose gradient multiplies the current to give the Joule heating shown in Figure 3-16.

The MATLAB code for this simulation is included in Section C.3.

We now double-check that the Peltier heat density generation rate shown in Figure 3-15 matches our expectations. The numerical integration of the Peltier heat density generation rate profile around one of the junctions reveals that  $Q_{\text{Peltier}} = 35.81\text{W}$ . Meanwhile, the difference in conduction band energy from the  $n^+$  to  $n$  region is  $35.82 \text{ meV}$ . Since the simulation was performed with a current of  $1 \text{ kA}$ , we would expect a heat generation rate of  $(1\text{e}3) \times (0.03582) = 35.82\text{W}$ . This result is in strong agreement with the numerically-integrated value, and the discrepancy can be explained by the distinction between the difference in transport energies and the difference in conduction band-edge energies across the junction as the density approaches the edge of the Boltzmann limit.

## 3.7 Summary and Conclusions

In this chapter we have developed a generalized drift-diffusion approach to microscopic thermoelectric transport and used it to describe both the Seebeck and Peltier effects. Using the momentum balance equation and Poisson's equation, we were able to simulate the Seebeck effect and achieve quantitative agreement while incorporating important aspects of modern thermoelectric materials, like doping beyond the Boltz-

mann limit, as well as details of current interest to the research community, namely inclusion of energy-dependent scattering times. Using the energy balance equation, we were able to express Peltier heating ( $Q_{\text{Peltier}} = \vec{J} \cdot \vec{E} - \vec{J} \cdot (-\nabla E_F)$ ) in terms of quantities available in any drift-diffusion simulation result.



# Chapter 4

## Applying the Framework

### 4.1 Chapter Overview

In Chapter 4 we seek to begin applying the theoretical framework derived from the reciprocal-space and real-space pictures of thermoelectric transport described in Chapter 2 and Chapter 3 respectively.

In Section 4.2 we expand on the discussion of the application space from Section 1.3, focusing in particular on the demands that the design processes for thermoelectric and micro-electronic devices would place on a simulation tool used in this capacity. We conclude that existing commercial software suits these needs, so in Section 4.3 we take up the task of replicating the Seebeck and Peltier effects as seen in Chapter 3. Next, in Section 4.4 we use the newly-validated software to examine the Peltier effect at a  $p$ - $n$  junction, and in Section 4.5 we briefly explore the limit of cooling based on the effect in the presence of Joule heating.

### 4.2 Application Space for the Theory

As we mentioned in Section 1.3, the purpose of developing microscopic models of thermoelectric phenomena is ultimately to generate improved designs for thermoelectric and micro-electronic devices.

For thermoelectric generators and coolers, recent research directions in the de-

velopment of more efficient modules warrant the use of computational models based on microscopic theory. Much of the recent progress made in improving  $ZT$  is owed to the introduction of nanoscale features which suppress thermal conductivity. Improvements in thermal conductivity have been shown based on nano-dot islands in a semiconductor matrix [7], meshes of rough silicon nanowires [27], and superlattices in the group V-IV alloys [28, 29]. While these nanoscale features benefit  $ZT$  by reducing  $\kappa_{ph}$ , they also introduce new boundaries which scatter charge carriers. As we saw in Section 3.2.3, whether the scattering dynamics of the carriers are dominated by these boundaries, isolated impurity atoms, or phonons may strongly affect the validity of the Relaxation Time Approximation and therefore all macroscopic calculations of thermoelectric efficiency that rely on it to accurately predict the Seebeck coefficient as well. In fact, looking directly to changing the dominating scattering process as a means of improving power factor has been recently proposed [24]. Clearly a need exists for a microscopic model like that derived from the Momentum Balance Equation in Chapter 3, which can both recover the results of the RTA and accurately describe the BTE in cases where it is not valid without resorting to computationally intensive methods like Monte Carlo.

Moreover, prospects of applying thermoelectric power generation to waste-heat recovery, including in automotive [21, 15] and industrial settings [15], have led to development of thin-film thermoelectric generators with low thermal resistance per unit area because the primary goal becomes fluxing a large amount of heat to maximize electrical power output rather than efficiency. These thin-film generators require the combination of various materials into a single leg because the range of operating temperatures is larger than the useful range of any individual TE material and avoiding thermal parasitics is key. Such devices are typically made by placing materials thermally and electrically in series (segmented) or in a combined series and parallel configuration (cascaded), and must be carefully designed to ensure thermoelectric compatibility [5, 6]. As we described in Section 1.2.2, the modeling for this design process is typically performed using the thermodynamic Onsager relations. However, because these thin-film generators have thicknesses on the order of microns,

it is computationally accessible to use a single microscopic model like the one presented in Chapter 3 to address the design of an entire leg. The present alternative is to embed the appropriate microscopic differential conductivity calculation, for which the relevant physics may change not only with material but temperature along the length of the leg, at each point in the macroscopic thermodynamic model and revisit the microscopic calculation iteratively.

Meanwhile, as described in Section 1.3, the operating characteristics of numerous micro-electronic devices depend strongly on the temperature at a particular point in the device. Instead of attempting to improve the larger systems they inhabit by cooling the entire system, it would be logical to consider Peltier cooling internal to a device in the design process. As we will see, the framework we presented in Chapter 3 is closely related to the models incorporated in commercial software packages for device simulation and design, making the inclusion of internally cooling in device design accessible.

By using such commercial software, we reap the benefits not only of using the highly-optimized numerical methods and huge databases of experimentally-verified material properties critical to the electronic device-design industry, but we also inherit a platform for compatibility. Moreover, many commercial device simulators also includes other relevant physics, including the joint solution of the full 3-dimensional heat transfer problem for the device. With this in mind, we seek to bridge the gap between the relatively simple computational models presented in Chapter 3 and commercial device simulators.

### **4.3 Transition to Commercial Device Simulators**

One popular commercial device simulation suite is Sentaurus Device, a Synopsys product. Sentaurus Device is used to perform 3-dimensional simulations for a range of micro-electronic [30] and opto-electronic [31, 32] device design applications, but it is capable of 1- and 2-dimensional simulations as well. While these options may become quite useful in the future, for example in accounting for thermal spreading

resistance as a thermal parasitic in a thermoelectric generator, for the problems that we will address in this chapter, simple 1-dimensional simulations will suffice.

Before we can make use of even a commercially-proven product to simulate thermoelectrics or micro-electronic devices that exploit thermoelectric effects, we must verify that we can replicate the results achieved in Chapter 3.

### 4.3.1 Replicating the Seebeck Effect

The Sentauros Device package contains four distinct transport models, which in order of increasing computational burden and physical accuracy are: the drift-diffusion model, the thermodynamic model, the hydrodynamic model, and Monte Carlo. Our hope is to replicate our physical model of the Seebeck effect with the simplest possible Sentauros transport model, so that when we turn to the problem of designing an inhomogeneous element for a thermoelectric generator, our simulations can converge on an optimized design on a reasonable timescale. The drift-diffusion model assumes the entire system to be at a fixed temperature, which is inconsistent with the boundary conditions of our Seebeck effect simulation, so the simplest available model is the thermodynamic model.

The thermodynamic model describes electronic transport with a handful of partial differential equations that are solved numerically. The PDEs are the Poisson equation, the continuity equations for electrons and holes, the heat equation for the lattice, and the following generalization of the drift-diffusion equation for each charge species:

$$\vec{J} = -nq\mu(\nabla E_F + \alpha\nabla T) \quad (4.1)$$

where  $\alpha$  is the Seebeck coefficient we hoped to calculate, as in Section 3.2. Sentauros offers an option to use the analytical expression for the Seebeck coefficient in the Boltzmann limit (1.33), letting the user adjust the scattering parameter, or amplify the entire coefficient by a multiplicative factor ( $\alpha' = m\alpha$ ).

There are a few reasons why this implementation may not be ideal for the devices we seek to simulate. First, having the user be given the option of amplifying the effect



offers an opportunity for the user to set a non-physical value of  $m \neq 1$ . Secondly, while (1.33) is only valid in the Boltzmann limit, as we saw in Section 1.2.3, the differential conductivity integral represents an inexpensive computational alternative to extend the result outside of the Boltzmann limit. As an alternative, the software could have chosen to calculate the Soret coefficient as we did in Chapter 2 and calculate currents in terms of gradients of electric potential, density, and temperature (as in the Momentum Balance Equation) instead of gradients in the electro-chemical potential and temperature (as in the differential-conductivity calculation). Incidentally, this is the approach taken in the more complex hydrodynamic model, which we will discuss shortly.

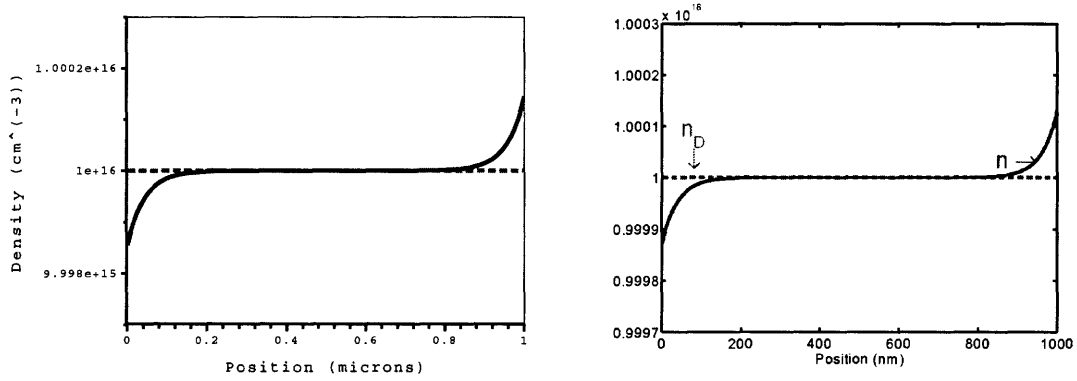


Figure 4-1: Left: carrier density profile for the Seebeck effect simulation in Sentaurus. Right: carrier density profile for the Seebeck effect simulation in generalized drift-diffusion code in Section C.2. Just as in the simulation from Section 3.2, the left boundary is fixed at 300.5 K and the right boundary is fixed at 299.5 K, leading to a temperature gradient of 1 K/ $\mu\text{m}$ . The remaining parameters used in both simulations were also identical, and can be found in the caption of Figure 3-2.

Nevertheless, we can use the thermodynamic model to validate our basic physical model for the Seebeck effect by performing a calculation at a doping concentration where the analytical solution remains valid. This is done for n-type GaAs with  $n_D = 1e16 \text{ cm}^{-3}$  at 300K, with the results as found in Figure 4-1 and Figure 4-2. To set up the temperature difference required in the Seebeck effect, we define so-called “thermodes” in Sentaurus to fix the edges of a semiconducting region to different temperatures. Sentaurus’s heat equation solver takes care of the rest, creating a

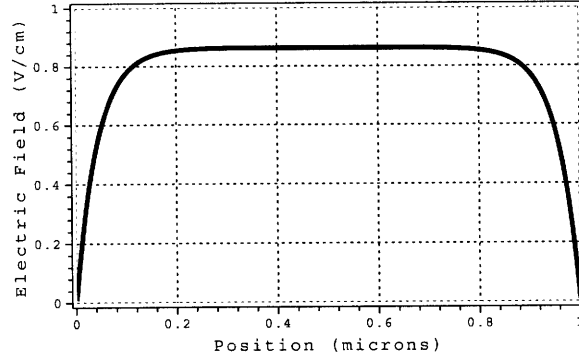


Figure 4-2: Electric field profile for the same Sentaurus Seebeck effect simulation as presented in Figure 4-1. We find that for energy-independent scattering times, the magnitude of the field away from the boundaries asymptotically reaches  $\approx 86 \mu\text{V/K} \times 1 \text{ K}/\mu\text{m} = 0.86 \text{ V/cm}$ .

roughly-linear temperature profile. The resulting density profile matches with that found in Section 3.2 and the electric field profile matches with that predicted by the Seebeck decomposition in Section 3.4.

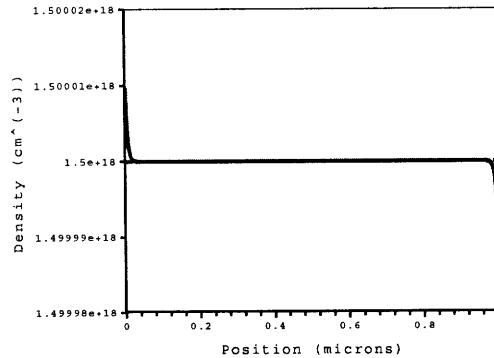


Figure 4-3: Carrier density profile resulting from invalid Sentaurus simulation of the Seebeck effect using the thermodynamic model. The ionized dopant concentration  $n_D = 1.5\text{e}18 \text{ cm}^{-3}$  corresponds to the doping level for peak power-factor, which is outside the Boltzmann limit. The software converges to a non-physical solution with free carriers piled up on the hot-side instead of the cold-side. All simulation parameters other than  $n_D$  are identical to those found in Figure 4-1.

Sentaurus Device does include a package to use Fermi-Dirac statistics when carrier densities are above the Boltzmann limit, where the thermoelectric power factor  $\alpha^2\sigma$  is maximized. However, in the thermodynamic model, this package serves only to permit carrier density calculations when the quasi-Fermi level is not well inside the bandgap.

It does not include a correction for the Seebeck coefficient, which in principle could be numerically approximated in a similar manner as the relationship between density and  $E_C - E_F$ .

If we ignore this issue and attempt to simulate the Seebeck effect for a doping concentration near the power-factor maximum as shown in Figure 3-11, which for GaAs was approximately  $n_D = 1.5 \times 10^{18} \text{ cm}^{-3}$ , the thermodynamic model will use the analytical solution outside its domain of validity to find an inappropriately-small  $|\nabla E_F|$ . Then, because the magnitude of the  $E_F$ -bandedge gap is set by material parameters and the temperature profile in the bulk region where the carrier density is approximately equal to the ionized dopant density, the results of the invalid Sentaurus simulation require  $\nabla E_C$  opposing  $\nabla E_F$ . Since the Poisson equation is also satisfied, the resulting non-physical density profile suggests that free carriers pile up on the hot-side of the device instead of the cold side, as shown in Figure 4-3.

To make use of Sentaurus to design thermoelectric elements, it appears we must use the hydrodynamic model, because it remains valid outside the Boltzmann limit, where as we saw in Section 3.5.1 the optimal- $ZT$  carrier concentration is found. In Sentaurus, the full hydrodynamic model seeks to numerically solve 8 coupled PDEs: the Poisson equation, the first three balance equations (the continuity equation, the momentum-balance equation, and the energy-balance equation closed by Fourier heat diffusion) for each carrier type and the heat equation for the lattice. The Fermi statistics package for the hydrodynamic model also modifies all of the transport coefficients, including the Soret coefficient in each momentum-balance equation, in principle permitting simulation of the useful TE generators. However, while the hydrodynamic model does work outside the Boltzmann limit, this improvement is accompanied by a host of more computationally-intensive features including gradients in the material's effective mass and distinction between the carrier and lattice temperatures. As a result, the complexity of setting up hydrodynamic simulations have impeded simulation of the Seebeck effect for power-factor optimizing dopant densities, leaving this task and the subsequent simulation of useful thermoelectric elements as future work (see Section 5.3).

### 4.3.2 Replicating the Peltier Effect

Since the Peltier effect is defined with an isothermal condition, it is accessible through the drift-diffusion model, the simplest transport model offered in Sentaurus Device. If we wish to accurately simulate a steady-state device with Peltier heating and cooling, we must use the more complex thermodynamic model which incorporates lattice and carrier heat transport, but for now the goal is to simulate the Peltier effect in an  $n$ - $n^+$ - $n$  homojunction device as we did in Section 3.6.

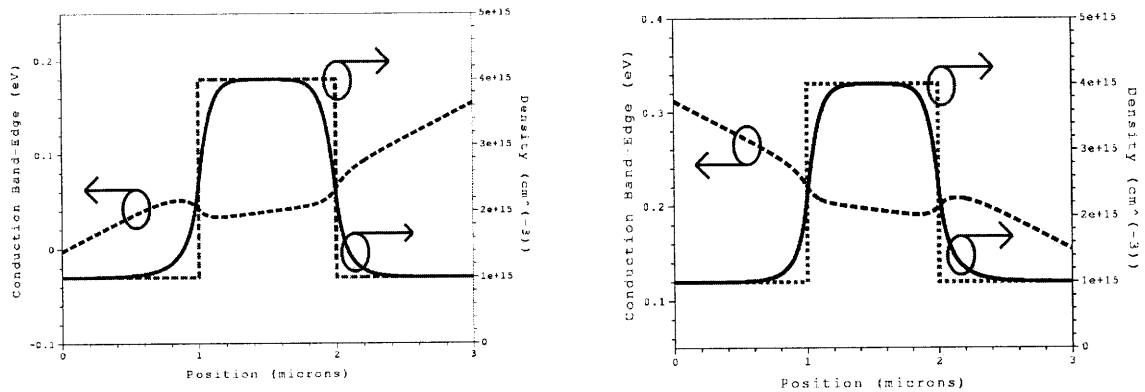


Figure 4-4: Results of driving a current of  $\pm 1$  kA/cm $^2$  through an  $n$ - $n^+$ - $n$  GaAs homojunction at 300K in Sentaurus. All other parameters are taken from GaAs, and are the same as those used in Figure 3-13. Left:  $J > 0$ . Right:  $J < 0$ . In each case, the conduction band-edge profile is plotted with units shown on the left, along with the donor-atom and carrier density profiles with units shown on the right.

We begin by generating an identical  $n$ - $n^+$ - $n$  homojunction structure to before, where the lightly-doped outer regions are given  $n_D = 1 \times 10^{15}$  cm $^{-3}$  and the more heavily doped inner region is given  $n_D = 4 \times 10^{15}$  cm $^{-3}$ . We then define two electrodes in Sentaurus and specify one to be a source terminal running a specific current, here  $\pm 1$  kA/cm $^2$ . This serves to current-bias the  $n$ - $n^+$ - $n$  homojunction device with  $\pm 1$  kA/cm $^2$ . The band-edge profile in Figure 4-4 matches the results from Figure 3-13. The ionized dopant and carrier density profiles also plotted in Figure 4-4 also match the results from the simulation in Section 3.6.2.

Sentaurus Device calculates various heat generation rates, which are consistent with the definitions presented in Section 3.6. The Joule heat generation rate for each

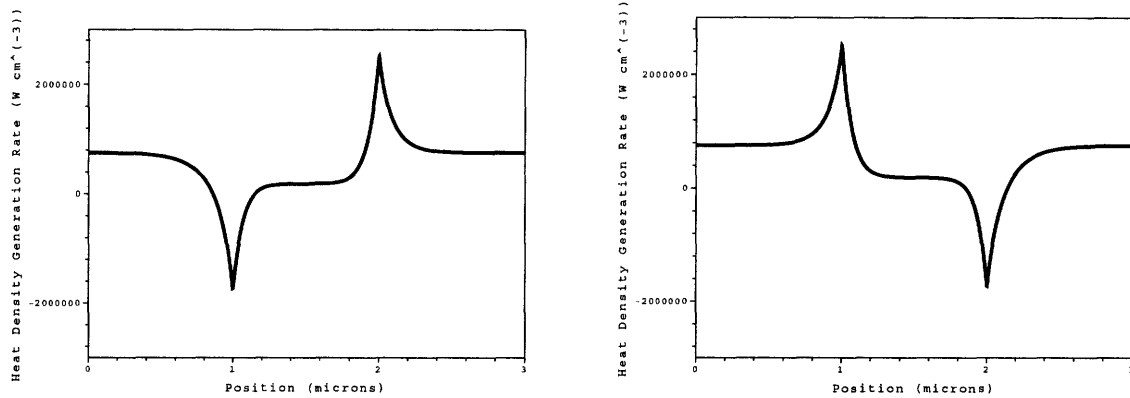


Figure 4-5: Total heat density generation rate ( $\vec{J} \cdot \vec{E}$ ) profile, where cooling appears as negative. Plot is for the same simulation as Figure 4-4, which uses Sentaurus to model the  $\pm 1 \text{ kA/cm}^2$  current-bias of an  $n\text{-}n^+\text{-}n$  homojunction.

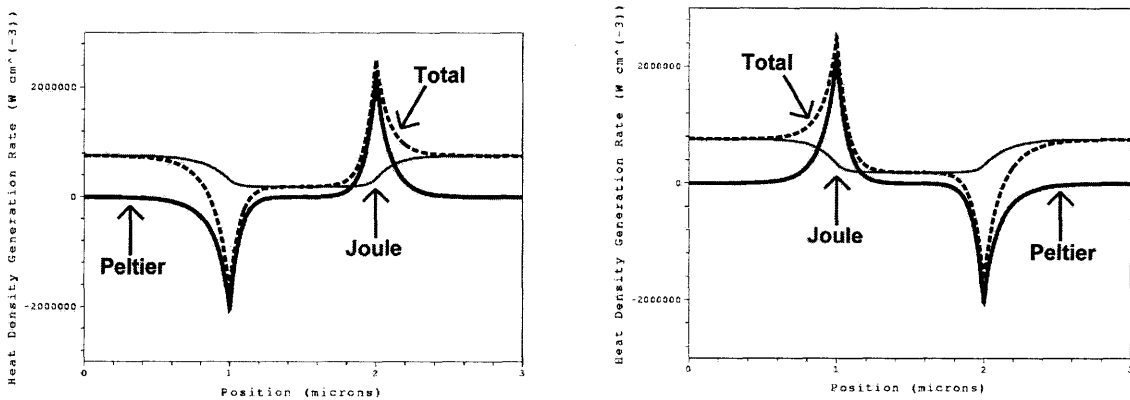


Figure 4-6: Plot showing decomposition of total heat density generation rate into Peltier and Joule terms. Again, cooling appears as negative heat generation. The thick solid line indicates Peltier heat; the thin solid line indicates Joule heat; the dashed line represents total heat, equal to the sum of the two. Plot is for the same simulation as Figure 4-4, which uses Sentaurus to model the  $\pm 1 \text{ kA/cm}^2$  current-bias of an  $n\text{-}n^+\text{-}n$  homojunction.

species is defined at each location by  $J^2/(qn\mu)$ . Since  $J/(qn\mu)$  has units of electric field, but net currents are only driven by gradients in  $E_F$ , we conclude that the Joule heat generation rate recovers our definition of  $\vec{J} \cdot (-\nabla E_F/q)$ . In Sentaurus, the Peltier heat generation rate for each species is defined at each location by  $(\vec{J} \cdot \nabla n) \frac{\partial \alpha}{\partial n}(-T)$ , which for an isothermal situation is equivalent to  $\vec{J} \cdot (-\nabla \Pi)$ . Interpreting the Peltier coefficient as the transport energy, and making use of the fact that the Seebeck coefficient was only properly defined in the Boltzmann limit where the transport energy and band-edge maintain a fixed separation, we can rewrite the Sentaurus definition of Peltier heat as  $\vec{J} \cdot \nabla(E_F - E_C)/q$ , which reconciles with the definition from Section 3.6. For the isothermal case here, where the Thomson effect can be neglected, the total heat generation rate is defined by Sentaurus to be the sum of the two, or  $\vec{J} \cdot (-\nabla E_C/q) = \vec{J} \cdot \vec{E}$ , also in agreement with Section 3.6.

The total heat generation profile is plotted in Figure 4-5 and is decomposed into its Peltier and Joule components in Figure 4-6. In Figure 4-5 as in Figure 3-14, the features of the heat generation near the junctions appear to reverse sign under current-reversal. Looking to Figure 4-6, we see that these features are in fact the reversible Peltier heat contributions. Note also that the Joule heating is greater in the regions with lower free carrier densities, as we would expect.

### 4.3.3 Translation of Sentaurus Variables

For the reader seeking to replicate the results presented here or otherwise use Sentaurus to model thermoelectric phenomena, a few details regarding the translation of quantities in Sentaurus into the language in this document may prove useful. Additionally, this process also provides some clarification of the results presented in Chapter 3, as it connects quantities like  $\vec{J} \cdot \vec{E}$  to more common terminology.

In all of our Sentaurus simulations we include dopants to create extrinsic semi-conducting regions. However, presumably because the primary market for the Sentaurus Device simulation suite is the silicon device-design community, the software's experiment-fitting database only includes silicon dopants. That is, while the dopant atoms themselves may be those we wish to implant in our GaAs substrate, the dopants

Sentaurus Variable	Interpretation (“BL” = Boltzmann Limit)
Joule Heat	$\sum_i \vec{J}_i \cdot (-\nabla E_{F,i}) \equiv \vec{J}_n \cdot (-\nabla E_{F,n}) + \vec{J}_p \cdot (-\nabla E_{F,p})$
Peltier Heat	$\sum_i (\vec{J}_i \cdot \nabla n_i) \left. \frac{\partial \alpha_i}{\partial n_i} \right _T (-T)$ where $\alpha_i$ is a function of $(n_i, T)$ . <ul style="list-style-type: none"> <li>· In BL (isothermal case), this is: <math>\sum_i \vec{J}_i \cdot (-\nabla \mu_i)</math>.</li> <li>· In BL (general case), this is: <math>\sum_i \vec{J}_i \cdot (-\nabla n_i \left. \frac{\partial \mu_i}{\partial n_i} \right _T)</math>.</li> <li>· Outside BL, this is non-physical as AnalyticTEP underestimates <math> \alpha_i(n_i) </math>.</li> </ul>
Thomson plus Peltier Heat	$\sum_i (\vec{J}_i \cdot \nabla \alpha_i) (-T)$ . <ul style="list-style-type: none"> <li>· In BL (isothermal case), this is: <math>\sum_i \vec{J}_i \cdot (-\nabla \mu_i)</math>.</li> <li>· In BL (general case), this is: <math>\sum_i \vec{J}_i \cdot ((-\nabla \mu_i) + \alpha_i \nabla T)</math> which is distinct from <math>\sum_i \vec{J}_i \cdot ((-\nabla \mu_i) + \left. \frac{\partial \mu_i}{\partial T} \right _{n_i} \nabla T)</math>.</li> <li>· Outside BL, this is non-physical as AnalyticTEP underestimates <math> \alpha_i(n_i) </math>.</li> </ul>
Recombination Heat	$R_{\text{net}}((E_{F,p} + T\alpha_p) - (E_{F,n} + T\alpha_n))$ where $R_{\text{net}}$ is the net recombination rate. <ul style="list-style-type: none"> <li>· In BL, this is: <math>R_{\text{net}}((E_{F,p} + \mu_p) - (E_{F,n} + \mu_n))</math>.</li> <li>· Outside BL, this is non-physical as AnalyticTEP underestimates <math> \alpha_i(n_i) </math>.</li> </ul>

Table 4.1: Translation of various heat-density generation rates calculated internally in the Sentaurus Device thermodynamic model into the notation of this document. Note that the electron chemical potential  $\mu_n$  is equivalent to  $E_F - E_C$ .

are defined by their properties in silicon (i.e. ionization energy, critical dopant-atom density above which complete ionization is assumed, etc.). While it is possible to implement user-defined species of dopants (this is in fact the procedure recommended by the Sentaurus manual for III-IV materials), a simpler solution is to simply use a common silicon donor (Phosphorous) and a common silicon acceptor (Boron). We have opted for this path in all III-IV simulations to present.

In the Seebeck effect thermodynamic-model simulation from Section 4.3.1, the thermo-electric power (another term for the Seebeck coefficient) is a quantity which

can be meaningfully defined for all semiconductors, but as we mentioned before the expression Sentaurus provides through its “AnalyticTEP” package is only valid in the Boltzmann limit. The overall scaling factor the package provides seems to the author to have no physical basis.

In the Peltier effect simulation from Section 4.3.2, where we used the simpler drift-diffusion model, Sentaurus calculates various heat density generation rates which correspond to the quantities in Table 4.1. We have attempted to express each heat-density generation rate in terms of a current density and a potential energy gradient. Although we avoid the issue in the simulations performed in this chapter, an intuitive interpretation of the “Recombination Heat” should be important when we consider long  $p$ - $n$  diodes (see Section 5.3) or attempt to reconcile our models with experimental data (see Section 5.2).

## 4.4 Internal Cooling at a $p$ - $n$ Junction

When a sample of semiconductor is  $p$ -doped in one region and  $n$ -doped in an adjacent region with a sharp junction between the two, the resulting structure is known as a  $p$ - $n$  junction and is a basic construct of micro-electronic devices. Because this junction rectifies currents flowing across the junction plane, when metal leads are attached to the ends of the  $p$ - and  $n$ -doped regions, the resulting device is referred to as a  $p$ - $n$  diode.

When a  $p$ - $n$  diode is at zero bias, no current flows because for both the electrons and holes individually, the drift and diffusion currents cancel exactly at every point in space. As the diode is put into forward bias, a net current flows as the diffusion current for each species overwhelms its diminished drift current. This net current is composed of a net flux of electrons moving to a region of lower electric potential ( $\vec{J} \cdot \vec{E} < 0$ ) and a net flux of holes moving to a region of higher electric potential (also  $\vec{J} \cdot \vec{E} < 0$ ). The quasi-Fermi levels of both the valence-band and conduction-band electrons are approximately flat through the depletion region over which the electric field is substantial, so in this region the contribution of Joule heating ( $\vec{J} \cdot (-\nabla E_F/q)$ )



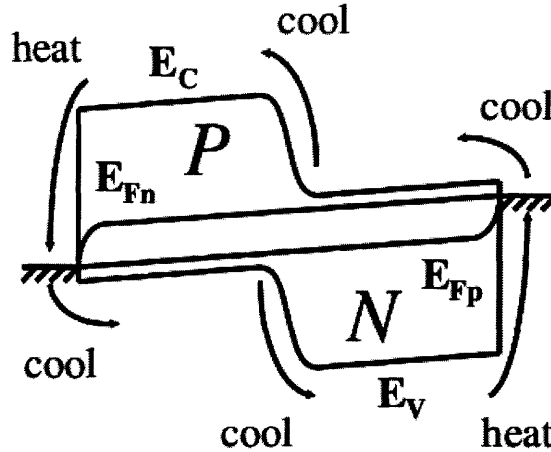


Figure 4-7: Depiction of Peltier heat exchange contributed by electrons and holes as a function of space. Since  $E_F$  for the two bands only re-equilibrate near the contacts, the picture takes the short-length approximation. Image from [2].

is negligible. As we derived in Section 3.6.1, we then expect that both species will reversibly draw heat out of the lattice and transport it through their minority carrier currents away from the junction to the edges of the device [2]. Depending on the equilibrium  $E_F$ -bandedge separation near the contacts there may also be Peltier heat exchange at the source contacts for each species. The locations at which the carriers in each band contribute to Peltier heating and cooling is shown in Figure 4-7.

Note that there is also Joule heating in the device, since the applied bias by definition requires the electro-chemical potentials within the contacts to differ by  $qV_{\text{applied}}$ . In the short-length approximation, where it is assumed that the only interaction between the electrons and holes besides Poisson's equation is recombination near the contacts, this Joule heating takes place in the contacts where the Peltier heat is also deposited by the minority carriers. In the long-length approximation, the  $p$ - and  $n$ -doped regions are much longer than the diffusion lengths so that recombination within the device re-equilibrates the quasi-Fermi levels for the two species in the bulk of the semiconductor, well before the contacts. On each side of the junction, the gradient of the minority-carrier  $E_F$  constitutes Joule heating, but since the bands remain approximately flat through this region, it would seem that our result from Section 3.6.1 implies that an identical amount of Peltier cooling takes place simul-

taneously. It seems that regardless of the nature of the irreversible recombination process, entropy is generated which binds some of the minority carriers' free energy into heat, which will be transported to the contacts where it will all be released as Peltier heat.

It is curious to note that the accounting of Peltier and Joule heat appears to be independent of the type of recombination process at play. For a direct-gap semiconductor the recombination could take place through emission of photons which leave the device, while in another material the emission of phonons that heat the lattice could be responsible.

What seems clear is that since the current carried by the minority carriers between the junction and the region of recombination is carried instead by the majority carriers between the region of recombination and the contact, we may ascribe this free-energy reduction to the flow of energy in the minority carriers, whose large chemical potential  $|\mu|$  indicates a large per-particle free energy, into the majority carriers whose small chemical potential indicates a small per-particle free energy.

A forward-biased  $p$ - $n$  diode was simulated in Sentaurus Device and the total heat density generation rate near the junction plotted in Figure 4-8. Since  $Q_{\text{Peltier}} \approx \vec{J} \cdot \vec{E}$  near the junction where  $\nabla E_{F,n}$  and  $\nabla E_{F,p}$  are negligible, we expect that the cooling should take place over the extent of the depletion region where the band-edges are not flat. The nearly-triangular shape of the cooling region shown in Figure 4-8 is indicative of the electric field profile from a nearly step-function charge density profile expected in the full-depletion approximation.

In the forward-biased diode simulation whose results are plotted in Figure 4-8, we note that the integral of the heat density generation rate over the  $1 \mu\text{m}$  device was  $-2.384 \text{ mW/cm}^2$  while the applied bias resulted in a total current density of  $\approx 4.888 \text{ mA/cm}^2$ . The ratio of these quantities give us the Peltier coefficient difference across the junction:

$$\Pi_n - \Pi_p = \frac{Q}{J} = \frac{-0.002384}{0.004888} \approx -0.4877\text{V} \quad (4.2)$$

This corresponds exactly with the observed  $\Delta E_C$  across the depletion region while

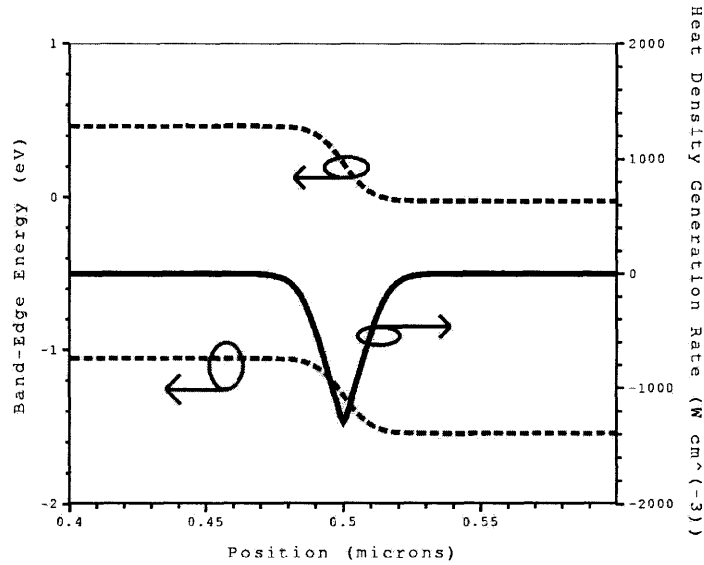


Figure 4-8: Profile of the heat density generation rate in a steady-state forward-biased GaAs  $p$ - $n$  diode. The solid line indicates the heat density generation rate with cooling shown as negative, and the corresponding scale is found on the right y-axis. The dashed lines indicate the profile of the conduction and valence band-edges measured relative to the equilibrium Fermi level at zero bias. The junction between  $p$ -doped region ( $n_A = 1e18$ ) and  $n$ -doped region ( $n_D = 1e18$ ) is modeled as abrupt, resulting in  $V_{\text{built-in}} \approx 1.49\text{V}$ . Diode as shown is forward biased at  $1\text{V}$ , resulting in a total current density of  $\approx 4.888 \text{ mA/cm}^2$ . Simulation taken with the length of the  $p$ - and  $n$ -doped regions as  $500\text{nm}$  each, though the region plotted above extends only  $100\text{nm}$  on each side of the junction. All material parameters for GaAs same as those taken in Figure 3-2.

under 1V bias, also found to be -0.4877V in the simulation. This confirms our intuition that the difference in Peltier coefficient corresponds to the difference in transport energy across the interface.

## 4.5 Maximal Internal Cooling in a Short $p$ - $n$ Diode

While the description of an operating  $p$ - $n$  diode from Section 4.4 indicates that Peltier cooling should be expected at the junction, we have thus far mostly neglected a thorough accounting of Joule heating. In diodes where the short-length approximation is valid, we might expect to see Joule heating near the contacts where equilibration between  $E_{F,n}$  and  $E_{F,p}$  takes place. In diodes where the long-length approximation is valid, the specific recombination process that dominates should influence the rate of heat-deposition into the lattice. For radiative recombination, the bulk of the power  $IV_{\text{applied}}$  could be absorbed by the contacts or leave the system entirely as light. For phonon emission pathways, this power should be absorbed by the lattice near the location of re-equilibration. Either way, the total Joule heating present throughout the device should be upper-bounded by energy conservation to be less than  $IV_{\text{applied}}$ . This quantity is represented visually alongside the rate of Peltier heat removal from the junction region in Figure 4-9.

However, regardless of the final destination of the energy from recombination, driving a sufficiently-large current should lead to Joule heating due to gradients in the majority-carrier  $E_F$  which are accompanied by gradients in the band-edges. This Joule heat should be expected intuitively, since the series resistance of a diode with sufficiently long  $p$  and  $n$  regions should result in sufficient Joule heating to match or exceed the Peltier cooling at the junction at high current, even in a short diode where all other heating takes place at the contacts. This Joule heating is proportional to the square of the current ( $Q_{\text{Joule}} \propto J^2 \rho$ ) while the Peltier cooling at the junction is to leading order linear in the current for small biases ( $Q_{\text{Peltier}} \propto \vec{J} \cdot \vec{E}_{\text{built-in}} \approx \vec{J} \cdot \vec{E}_{\text{built-in}}|_{J=0}$ ) and thereafter sub-linear. Thus, even neglecting any Peltier and Joule heating in the contacts on the grounds that it could be efficiently extracted to

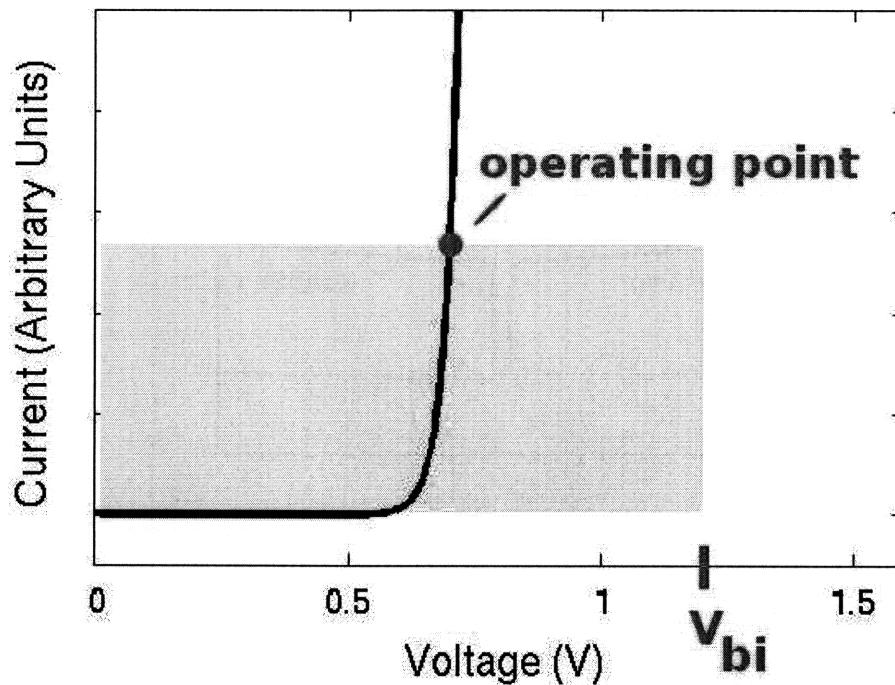


Figure 4-9: Diagrammatic description of heat generation rates in a forward-biased short-length  $p-n$  diode. The area corresponding to  $IV$  (left block) represents the total Joule heat throughout the device and the area corresponding to  $I(V - V_{\text{built-in}})$  (right block) represents the Peltier heat which is transported from the junction to the contacts.

the environment, over the extent of the semiconducting region of the device, applying a small bias to drive a small current will result in net cooling but applying enough bias to drive a sufficiently large current will result in net heating.

For the diode examined in Section 4.4, applying a forward bias of 1.5V results in substantial heating throughout the bulk of the device, as seen in Figure 4-10. Above a bias of roughly 1.488V, the net heating in the device is positive as the Joule heating contribution throughout the  $1 \mu\text{m}$  body overcomes the cooling at the junction, as seen in Figure 4-11. Although the extremely large current densities required (hundreds of  $\text{kA}/\text{cm}^2$ ) indicate that the assumption of infinite heat sinking at the contacts is invalid, the analysis nonetheless demonstrates the existence of a bias-point for maximum net cooling in even a short  $p-n$  diode.

Although the long-length case was not explored computationally, Sentaurus does

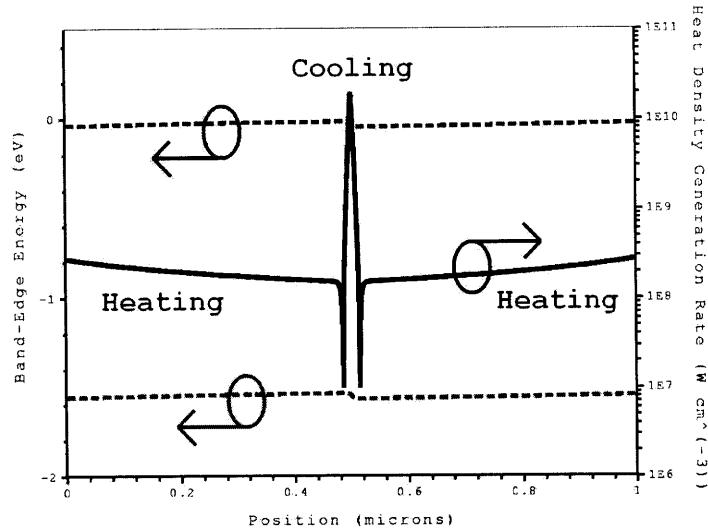


Figure 4-10: Heat density generation rate profile for the same device as in Figure 4-8, but with a 1.5V bias applied instead of 1V, resulting in a current density of 444.0 kA/cm<sup>2</sup> instead of 4.888 mA/cm<sup>2</sup>. This plot is for the entire device, not just a small region surrounding the junction as in Figure 4-8. The solid line indicates magnitude of heat density generation rate, plotted on a log scale (right y-axis) so that cooling and heating both appear positively. The region near the junction experiences net cooling, but beyond the minima of heat generation rate magnitude the device experiences net heating. The dashed lines again indicate the band-edge profiles (left y-axis).

model the necessary physical processes to permit the re-equilibration of  $E_{F,n}$  and  $E_{F,p}$  away from the device's contacts. Interestingly, while the long diode may experience more Joule heating, we may see a lower minimum temperature at the junction than in a short diode. To see why, consider that at sufficiently small bias, the ratio of Peltier to Joule heat is large, but if the small current is extracting Peltier heat from the junction and depositing it at the contacts at a given rate, these source terms in the heat equation will lead to a specific temperature gradient between each contact and the junction. Although this gradient will need to be smaller at the  $\Delta T$ -optimizing current than in the short-length case to suppress Joule heating, if the device is longer the temperature difference may be larger. We hope to explore this problem further computationally.

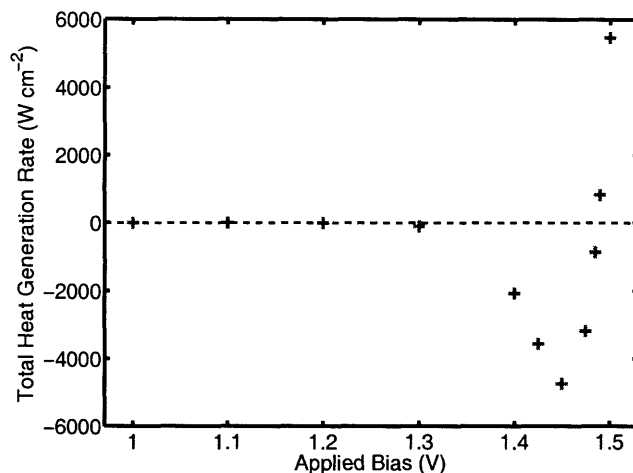


Figure 4-11: Plot of total heat generation rate, given by integrating the total heat density generation rate over the  $1 \mu\text{m}$  length of the  $p$ - $n$  diode simulated in Figure 4-8, versus bias voltage. For sufficiently large bias, there is net heating. Note that this calculation omits heating at the contacts, permitting net cooling for  $V_{\text{applied}}$  below  $\approx 1.488\text{V}$  or  $J_{\text{driven}}$  below  $\approx 342 \text{ kA/cm}^2$ .

## 4.6 Summary and Conclusions

The modeling of thermoelectric effects based on the microscopic semi-classical transport theory presented in Chapter 2 and Chapter 3 has been replicated in the commercial device physics simulator Sentaurus Device. The use of such simulators makes accessible more realistic simulations, both by inclusion of a wider array of physical effects and by implementation of superior numerical algorithms, and offers a platform for the development of improved designs for thermoelectric generators and coolers as well as the exploration of thermoelectric cooling in micro-electronic devices with the goal of designing internally-cooled devices. The origin of Peltier cooling at a  $p$ - $n$  junction has been explained, and the limitation of Joule heating has been explored in a simple  $p$ - $n$  diode, though quantitative reconciliation of these theoretical results with initial experiments remains future work (see Section 5.2). Since  $p$ - $n$  junctions represent basic constructs from which many micro-electronic devices are built, these results may directly aid in the development of internally-cooled devices, including HBTs, BJTs, and diode lasers.





# Chapter 5

## Summary and Future Work

### 5.1 Summary and Chapter Overview

In Chapter 1 we explained that the ultimate goal of this research was to develop improved designs for inhomogeneous thermoelectric elements and micro-electronic devices.

In Chapter 2 we developed a reciprocal-space picture of the fundamental transport processes responsible for the thermoelectric effects, and in Chapter 3 we used these results to explain the Seebeck and Peltier effects in a generalized drift-diffusion framework. These results constitute the conceptual basis for the simulations done with Sentauros in Chapter 4, including those demonstrating the limits to Sentauros's models and those showing internal cooling in a basic  $p$ - $n$  diode.

Although the framework has been built, much work remains before novel designs of inhomogeneous thermoelectric elements and internally cooled micro-electronic devices can be realized. To that end, in Section 5.2 we outline the path to experimental verification of our microscopic model of thermoelectricity and propose two experiments. In Section 5.3, we discuss the physical effects which must be incorporated into simulations to enable realistic design and mention some generally promising directions on the theoretical front.

## 5.2 Future Experimental Work

### 5.2.1 Experimental Verification of Internal Cooling Model

Though the author has not contributed to the ongoing experimental work described in this section, initial comparisons with the preliminary (unpublished) results represent a unique opportunity to verify the theoretical results presented in this work. We briefly describe the experimental results achieved by the author's collaborators before presenting the results of a Sentaurus simulation of this experiment and speculating on the discrepancies.

Sample #1	Sample #2		
$N_A=1 \times 10^{20}$	$N_A=1 \times 10^{20}$	p+ Cap	0.1 $\mu\text{m}$ $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$
$N_A=1 \times 10^{19}$	$N_A=2 \times 10^{18}$	p	1.0 $\mu\text{m}$ $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$
$N_D=1 \times 10^{19}$	$N_D=2 \times 10^{18}$	n	1.0 $\mu\text{m}$ $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$
$N_D=3 \times 10^{18}$	$N_D=3 \times 10^{18}$	Substrate	InP

Figure 5-1: Layer structure of InGaAs diode on which initial junction cooling experiments were done. Doping concentrations in  $\text{cm}^{-3}$ . Image from [2].

A short  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$   $p$ - $n$  diode, depicted in Figure 5-1, was grown on an InP wafer by Shaomin Wu and Prof. John Bowers at UC-Santa Barbara. Next, a small-area top contact and a large-area contact beneath the substrate were deposited by Kevin Pipe. The sample was then current-biased between these contacts, and its temperature profile was imaged using stochastic resonance-enhanced thermoreflectance by Dietrich Lürßen. The temperature profiling results are conveyed in Figure 5-2. The temperature of the  $p$ - $n$  junction is lower than that of the substrate in forward bias and the magnitude of the temperature difference increases with current. The data set does not extend to larger current densities because such conditions resulted in destruction of the device.

We simulated this structure in Sentaurus using a 1-dimensional thermodynamic model with full Fermi-Dirac statistics. Because the doping levels in the  $p$  and  $n$

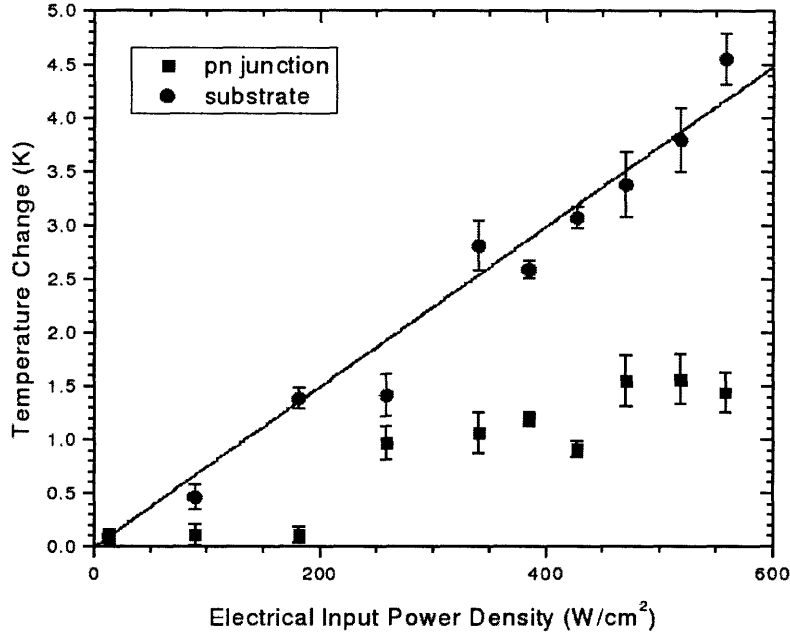


Figure 5-2: Dependence of temperature on input power density ( $I_{\text{driven}} \times V_{\text{applied}}$ ) at junction (squares) and  $2 \mu\text{m}$  into the substrate (circles). Image courtesy of Dietrich Lürßen and Rajeev Ram.

regions make both degenerate, we also enabled local band-to-band tunneling, though the operating bias of  $0.6\text{V}$  suppressed all tunneling current flow for the situations described here. We found that the  $0.6\text{V}$  forward bias, which produced a current density of  $\approx 1.055 \text{ kA/cm}^2$  (see Figure 5-4), corresponded to an input power density of  $\approx 630 \text{ W/cm}^2$ . According to the experimental data in Figure 5-2, we would expect to see cooling at the junction in the range of  $3\text{K}$  compared with the substrate region. Instead, we see cooling of just  $0.2\text{K}$ . There are many possible explanations for this discrepancy.

First, looking at the temperature profile in Figure 5-4, we see that the thermal boundary conditions employed are problematic. The temperature is fixed to be  $300\text{K}$  at a point just  $2 \mu\text{m}$  below the InGaAs  $n$ -region where we would expect the temperature to be above ambient because of Joule heating, as seen in Figure 5-2. In an initial exploration of this possibility, it was found that extending the  $2 \mu\text{m}$  substrate region to  $\approx 300 \mu\text{m}$  resulted in the minimum temperature at the junction being reduced to  $299.75\text{K}$  from  $299.80\text{K}$ , but was accompanied by a decrease in the temperature  $2 \mu\text{m}$

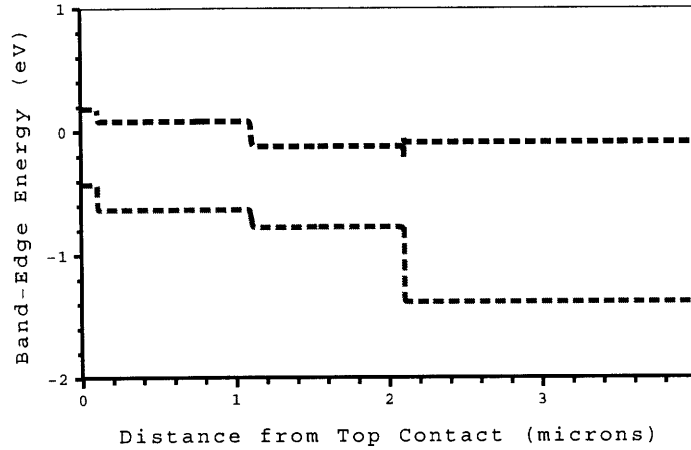


Figure 5-3: Band-edge profile for simulated structure under 0.6V, 1.055 kA/cm<sup>2</sup> bias. The left edge corresponds to the top contact and the right edge corresponds to a point 2  $\mu$ m into the substrate

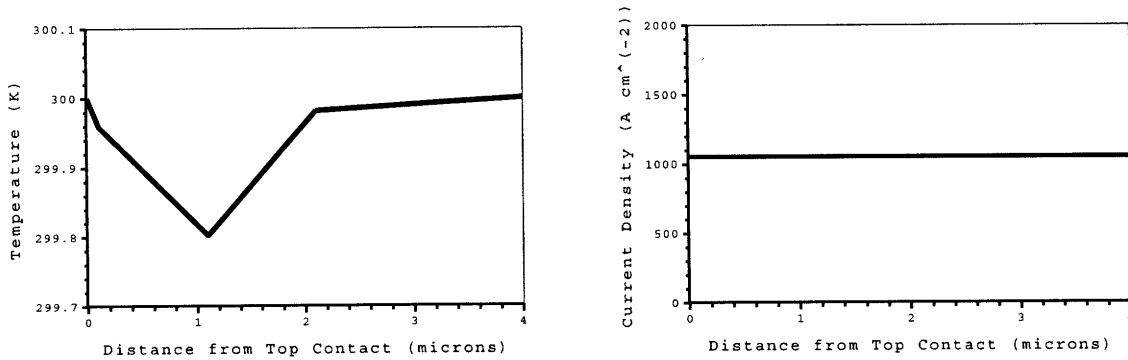


Figure 5-4: Left: Sentaurus simulation results for temperature profile of InGaAs diode structure under 0.6V bias. Right: current density profile. This current density corresponds to  $\approx 630$  W/cm<sup>2</sup>.

into the substrate to 299.80K from 300K. That is, the longer substrate simulation further reduced the  $\Delta T$  due to cooling from 0.2K to 0.05K. Modifying the thermal boundary condition to have a fixed heat-transfer coefficient instead of a fixed temperature would be more physical, and may prove a more effective approach as it should permit temperatures well above ambient to persist in the device.

Next, the simulation conducted is a 1-dimensional simulation, but the small top contact and large substrate area suggest that substantial thermal and electrical spreading should be present during device operation. Sentaurus is capable of higher-dimensional simulations, even though they were not initially selected.

Both the thermal boundary-condition problem and the multi-dimensional transport problem can in principle be addressed with Sentaurus, and their exploration as possible sources of discrepancy with the experimental data are left as future work.

Though the counter-intuitive simulation results suggest greater problems with the numerical model, in order to achieve agreement we must also re-examine our physical model of the experiment and the experimental results themselves. We therefore note that since thermoreflectance imaging requires a cleaved boundary for light to reflect off or be absorbed by, the experimental procedure is not really measuring the temperature profile within the bulk of the device but rather near the surface. As a result, boundary scattering of phonons near this surface may be lowering the effective thermal conductivity near the surface, which could lead to a larger steady-state temperature gradient for the same cooling power, resulting in a larger temperature difference between the junction and the nearby substrate. Finally, errors in experimental interpretation must also be considered.

### 5.2.2 Experimental Proposals

The theoretical description of the thermoelectric effects expressed in this document have given rise to two new experimental proposals.

The setup of the first experimental proposal is depicted in Figure 5-5 and the expected measurement is depicted in Figure 5-6. Our microscopic real-space picture of the Seebeck effect suggests that the gradient in the vacuum level, reflected in the local electric potential gradient near the surface of a thermoelectric element in a Seebeck effect setup, should not correspond to the full thermoelectric voltage divided by the sample length. Instead, as we saw in Section 3.4, we expect the longitudinal electric field to be proportional to the temperature gradient present with a constant of proportionality equal to  $(k_B/q) \times (1+p/2)$  where as usual  $p$  indicates the power-law dependence of the carrier relaxation time on wave-vector ( $\tau(\vec{k}) \propto |\vec{k}|^p$ ).

This may be of some use as this experiment could provide an independent measurement of the scattering parameter governing the dynamics of the carriers in the Boltzmann limit. If the responsible scattering processes continue to dominate at

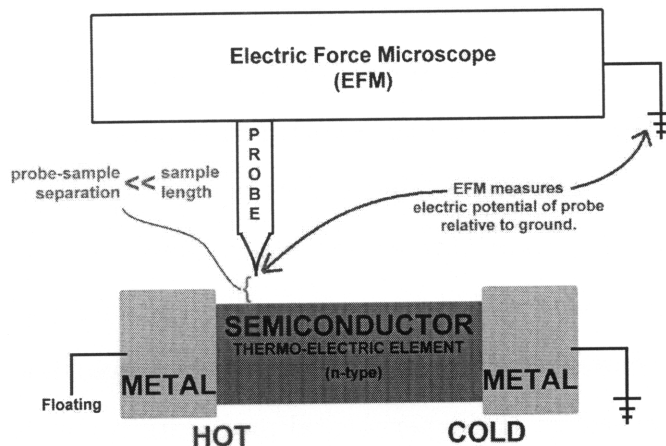


Figure 5-5: Proposed setup for an experiment to measure the scattering parameter governing electron dynamics in an  $n$ -type thermoelectric material. The material would need to be doped weakly to remain in the Boltzmann regime, but the results may remain valid at the higher doping levels where  $ZT$  is maximized.

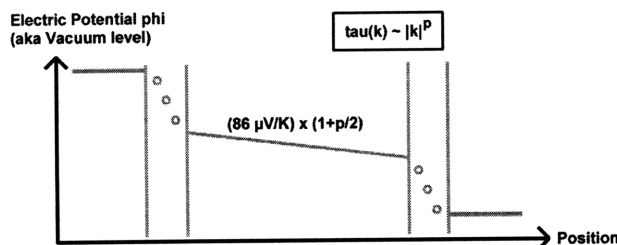


Figure 5-6: Depiction of the expected measurement from the scattering-parameter experiment depicted in Figure 5-5.

higher doping concentrations, this may be useful [24] for the development of thermoelectric materials with higher power factors and improved  $ZT$ s. However, the experiment should serve mostly as a validation of the Seebeck effect model presented in Chapter 3.

A second experimental proposal that has grown out of the theory presented in this document is for the construction of a 3-terminal thermoelectric device, in which two of the terminals serve as a source and drain to drive current through a thin slab of semiconductor (or perhaps even a low-dimensional electron gas). The carriers should be confined to a thickness shorter than the Debye screening length in the material.

We then attach an electrode (the gate) above the surface of the semiconductor in such a configuration as to have the current path lead from the source, through a region not near the gate, then through the region under the gate, then out to a region not near the gate before reaching the drain. Since the tight capacitive coupling between the part of the current path under the gate and the gate electrode would permit the gate electrode to deplete or exceed the quasi-neutral carrier concentration and locally modifying the profile of the band-edge. As a result, the gate would control the Peltier heating and cooling which would take place at the intersection of the current path and the edge of the semiconducting region in close proximity to the gate. Positive gate voltages would cause heating at one end and cooling at the other, while negative gate voltages would cause the reverse.

### 5.3 Future Theoretical Work

In spite of the basic framework put in place in this document, several theoretical challenges remain.

First, although we have simulated the Seebeck and Peltier effects and explained that Sentaurus in principle contains the physics required to accurately simulate and design complete TE coolers and generators, the task of doing so has yet to be done.

In order to conduct simulations useful for the design of thermoelectric elements, we must be able to simulate the materials of interest. As mentioned in Section 3.5.2, the bandstructures of most useful materials to date have included numerous degenerate ellipsoidal valleys away from the zone-center  $\Gamma$ . Resolving the correct way to use the results of Section 2.3 to calculate transport coefficients and then use the tools from Chapter 3 to simulate inhomogeneous thermoelectric elements currently under research remains a high priority going forward. Such investigations are likely to involve the simulation and design of elements incorporating functional grading and segmentation; the design of cascaded thermoelectric elements may also be enabled by the 2- and 3-dimensional heat-transfer solvers built into Sentaurus.

On a more theoretical level, the results explaining the Peltier effect in a drift-

diffusion framework and their application to even simple micro-electronic devices retain a bit of mystery. The complete accounting of the entropy and energy flows among the various subsystems (electronic carriers, the lattice, and the free-space electromagnetic modes) in any real-life example of the Peltier effect is rather complicated and somewhat counter-intuitive. As an example, consider an LED forward biased below turn-on. At such small biases, the current which flows is nonzero and the electrical power input to the device is  $IV_{\text{applied}}$ . Nevertheless, each electron which contributes to current must undergo a local radiative recombination event and emit a photon of energy  $E_{\text{gap}}$ . If such photons are efficiently coupled out of the device, the energy emitted in photons is greater than the electrical energy input by the ratio  $(E_{\text{gap}}/q)/V_{\text{applied}}$ . The perspective of the picture of the  $p$ - $n$  diode from Chapter 4 suggests that the cooling at the junction should exactly make up the difference, as the ratio of the Peltier power to input power is as indicated by Figure 4-9.

The picture we are left with is that of an extremely efficient heat pump, operating between a cold-side given by the lattice temperature and a hot-side determined by the emitted photons. Here we see a common trade-off in non-ideal heat engines between power and efficiency. If we wish to pump energy from the lattice to the free-space electromagnetic modes efficiently, we must do so at low bias where current is low. If we wish to pump energy at high power, a dominant fraction of the energy must eventually come from the electrical power source and not the lattice.

Although the author has not carefully performed the calculation, it is interesting to note that the entropy per unit energy in a light source is dependent on the spectrum alone. On this basis, we may expect that the sub-turn-on forward-biased LED may act as a heat engine operating between one reservoir at the temperature of the sample and another at a temperature determined by the output spectrum.

In addition to the purely theoretical curiosities suggested here, some more practical questions remain unanswered. For example, most applications of the Peltier effect make use of junctions between different material systems, including those between metal and semiconductor in thermoelectric power conversion applications or at a semiconductor hetero-junction in an internal cooling device application [2]. Nevertheless,



we have not yet extended our microscopic description from Section 3.6 to include such junctions. We note that while the general approach outlined in Section 3.6.1 is applicable to the case of heterojunctions, a range of material properties change across this interface and not just the conduction-band ionization energy  $E_{vac} - E_C$ . For primarily this reason, we have left the simulation of the Peltier effect at heterojunctions as future work.

In order to match experimental data and to understand short devices, we must also more effectively include the effects of the contacts. These should include the Joule heating due to contact resistances and the Peltier heat exchange we have discussed at length, as well as the effects of finite heat-transfer rates to the environment. We note that Sentaurus does include the physics necessary to model these effects.

Finally, this perspective on device physics offers possibilities for the improved design of a wide range of micro-electronic devices whose operation is degraded by heating. The understanding taken from the connection between microscopic transport and Peltier cooling established in Section 3.6 and its specific application to a basic building block of micro-electronic devices in Chapter 4 may lead to the development of improved designs for devices, including HBTs and diode lasers.



# Appendix A

## The Semi-Classical Picture

### A.1 The Boltzmann-Transport Equation

The Boltzmann Transport Equation (BTE) is the fundamental equation of non-equilibrium statistical mechanics and semi-classical electron dynamics. It can be used to describe the dynamics of a population of weakly-interacting particles in free space, and with the benefit of Bloch's theorem and bandstructure, weakly-interacting electrons in a periodic lattice such as a solid.

The BTE is unique in that depending on how scattering is included, it can be used to describe purely reversible or irreversible dynamics. The BTE may be derived and a rigorous meaning may be given to the scattering term responsible for irreversibility through the BBGKY Hierarchy (see Section A.2). For simplicity we begin by taking as Boltzmann did [?] the *Stosszahlansatz*, which states that the distribution functions for each particle are uncorrelated so that we may introduce the effects of scattering based on statistical rather than purely mechanical observations. Upon doing so, we arrive at the BTE as it is typically presented:

$$\frac{d}{dt}f(x, k, t) = \left[ \frac{\partial}{\partial t} + \vec{x} \cdot \nabla_x + \vec{k} \cdot \nabla_k \right] f(x, k, t) = \hat{S}[f(x, k, t)]$$

Let us begin by describing the central object in the semi-classical BTE transport formalism,  $f(x, k, t)$ . This object is often referred to as the one-particle phase-space

density or simply the distribution function. It is possible (because we assume ergodicity) to interpret it either as the probability density for any one of the  $N$  particles to be found at a given point in phase-space or alternatively as the fraction of the  $N$  particles which are found there at time  $t$ . Let us imagine it to represent the density of actual particles for the time being, as it makes the picture behind the BTE more physical.

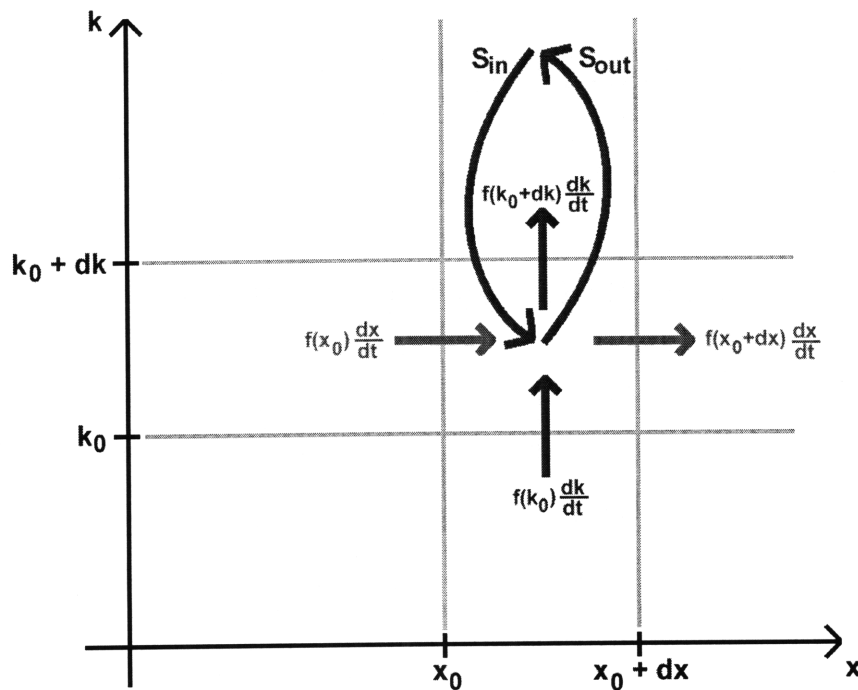


Figure A-1: Graphical representation of the phase-space density flow described by the Boltzmann Transport Equation (BTE).

Let us imagine looking closely at the distribution function  $f$  in a small region of the phase-space of independent variables  $x$  (position) and  $k$  (crystal momentum). The quantity  $f(x_0, k_0, t) \cdot dxdk$  describes the number of particles which can be found within a region of width  $dx$  around  $x_0$  and within a region of width  $dk$  around  $k_0$ . For a finite concentration of particles in both configuration and momentum-space to be consistent with the Heisenberg uncertainty principle, we must consider them to inhabit wavepacket states with position and momentum centered around the point under consideration. For our purposes, it will suffice to keep this fact out of mind, and to proceed as though both quantities could be simultaneously well-defined for

each particle. This assumption constitutes the essence of the semi-classical picture.

Derivations of bandstructure and semi-classical equations of motion for electrons are presented in many introductory solid-state textbooks [?] and we will not repeat them here. Instead, we simply recall that the semi-classical equations of motion tell us that the rate of change in position of a particle is proportional to the crystal momentum ( $\dot{x} = \frac{\hbar}{m^*}k$ ) and that the rate of change of crystal momentum is proportional to the external force applied to it ( $\dot{k} = \frac{1}{\hbar}F_{ext}$ ). Given these equations, and supposing the existence of a scattering function  $S(k, k')$  describing the average rate for electrons with crystal momentum  $k$  to make a transition to  $k'$  according to a Poisson process, we account for all of the arrows in the figure above and relate them to terms in the BTE.

## A.2 The BBGKY Heirarchy

This framework permits one to describe the dynamics of a one-particle distribution function (i.e., the object  $f$  which we will soon describe) in terms of a two-particle distribution function, and the dynamics of a two-particle distribution function in terms of a three-particle, and so on. Ultimately the dynamics of the one-particle distribution function depend on the individual reversible dynamics of the particles expressed through the  $N$ -particle distribution function, where  $N$  is the total number of particles in the system.

$$\begin{aligned}
 \hat{D}[f(x, k, t)] &= \hat{I}[f_2(x_1, k_1, x_2, k_2, t)] \\
 \hat{D}[f_2(x_1, k_1, x_2, k_2, t)] &= \hat{I}[f_3(x_1, \dots, k_3, t)] \\
 &\vdots \\
 \hat{D}[f_{N-1}(x_1, \dots, k_{N-1}, t)] &= \hat{I}[f_N(x_1, \dots, k_N, t)]
 \end{aligned}$$

However, working analytically or computationally with a function whose indepen-

dent variables number at the Avagadro-scale is intractable, so we typically choose to cut off this Heirarchy at the one-particle distribution function and replace its dependence on the two-particle distribution function with a statistical (but heuristic from the point of view of reversible dynamics) functional to describe the influence of scattering.

$$\hat{D}[f(x, k, t)] = \hat{S}[f(x, k, t)]$$

It is this functional  $\hat{S}[f]$  which is responsible for the resulting irreversible dynamics. While the BBGKY may begin to clarify the question of how irreversible dynamics result from a statistical composition of the reversible microscopic laws which govern the individual particles, it does not offer a particularly intuitive or applicable picture, and hence we have chosen to motivate the equation more heuristically in Section A.1.

# Appendix B

## Non-Equilibrium Thermodynamics of Thermoelectricity

In this section we intend to construct an explicit thermodynamic framework which appropriately describes transport in systems near equilibrium. Much of this discussion follows Chapter 14 of the text by Callen [3].

Systems out of thermodynamic equilibrium (whether in steady-state or not) do not fall within the domain of validity for results from equilibrium thermodynamics. Nevertheless, much can be learned about systems near equilibrium from the approach to equilibrium as presented in traditional thermodynamics.

Consider, for example, a pair of identical physical systems, but with one at temperature  $T_A$  and the other at  $T_B < T_A$ . When the systems are brought into thermal contact with one another but are kept adiabatically insulated from the rest of the universe, energy will flow from A to B to increase entropy (or decrease free energy) of the composite system until:

$$\frac{\partial S}{\partial U_A} = \frac{\partial(S_A + S_B)}{\partial U_A} = \frac{\partial S_A}{\partial U_A} - \frac{\partial S_B}{\partial U_B} = 0$$

since the total energy  $U = U_A + U_B$  is fixed. Since temperature is defined as  $T \equiv (\frac{\partial S}{\partial U})^{-1}$ , we see this as the thermal equilibrium condition as expected:  $T_A = T_B$ . We may interpret this process as the system traversing a path in thermodynamic

state-function space which represents the steepest ascent of entropy. The “flux” of energy flows from system A to system B until the “force” of the temperature difference disappears, rendering a combined system in which there is no entropy-increase to be gained by the transfer of energy.

Similarly, we may define other “forces” and “fluxes” which bear the relationship to one another, just as the inverse temperature difference and energy did in the example above, that the flux is zero when the force vanishes, and for small values of the force (corresponding to states of the system close to thermodynamic equilibrium) the value of the flux is linear in its driving force. Hence, as we see in the example above, the key to connecting forces with their corresponding fluxes is the potential to generate entropy. In fact, for systems where the exchange of multiple extensive quantities is permitted, we can assign an origin to each component of the entropy generation:

$$\dot{S} = \sum_k \frac{\partial S}{\partial X_k} \frac{\partial X_k}{\partial t}$$

The entropy generation rate is the product of each extensive quantity flow rate (flux) with its associated means of generating entropy (force).

## B.1 Continuous Systems

An analogous relationship can be defined for continuous systems, where entropy density replaces entropy and the fluxes become three-dimensional flow densities (extensive quantity per unit cross-sectional area per unit time). In this picture instead of the differences in intensive quantities that we called “forces,” we refer to gradients of intensive quantities called “affinities,” which indicate the direction of flux which corresponds to the maximal rate of local entropy density generation. We define for each flux a corresponding affinity whose contours define surfaces along which fluxes do not change the local entropy density.

This point can be somewhat confusing because the overlapping language with the discrete case above belies an additional complexity. The definition above neither



precludes nor enforces the flux in the directions along the surface to be zero; it merely states that any such flux will not contribute to the local entropy density generation via the affinity in question. In the discrete case above, our “flux” referred to a rate of extensive quantity flowing from one specified subsystem to another. For the continuous 3-dimensional case, for the sake of interpreting the definition above, the analogous quantity is not the typical flux vector (vector quantity per unit time per unit cross-sectional area) but the component of the flux vector in the direction of the affinity gradient. For the case of energy flow and temperature, the component of energy flux pointing up the inverse temperature gradient must be zero when and only when that inverse temperature gradient disappears. The flow of energy along surfaces of constant inverse temperature is not necessarily zero, nor does its value tell us anything about the inverse temperature gradient present in the system. Keeping this fact in mind, we will still express our results algebraically in terms of the familiar flux vectors to keep the notation compact.

Since in the continuous case, such as the flow of heat down a bar with a temperature gradient across it, no finite volume is truly in equilibrium, we choose to examine the system in terms of differential volume elements. Each volume element is presumed to be in equilibrium itself, but at a different point in state-function space than its neighbors. Nevertheless, we assume that all of these infinitesimal subsystems have the same relationships between their extensive quantities as the macroscopic system would if it were in equilibrium. That is, the fundamental equation (working in the entropy representation):

$$dS = \sum_k F_k dX_k$$

tells us that the entropy per unit volume ( $s$ ) obeys:

$$ds = \sum_k F_k dx_k$$

Now that we have an expression for an infinitesimal unit of entropy density, we may express the flow of entropy in terms of the flow of other extensive quantities simply

as:

$$\vec{J}_S = \sum_k F_k \vec{J}_k$$

In the case of thermoelectricity, the only fluxes of interest are the flux of energy  $\vec{J}_U$  and of particles  $\vec{J}_N$ :

$$\vec{J}_S = \frac{1}{T} \vec{J}_U - \frac{\mu_{ec}}{T} \vec{J}_N$$

where  $\mu_{ec}$  is the electro-chemical potential, referred to elsewhere as  $E_F$ . Now, the continuity of entropy flow tells us that the rate of entropy density generation is equal to the sum of its accumulation and outflow:

$$\dot{s} = \frac{\partial s}{\partial t} + \nabla \cdot \vec{J}_S$$

Meanwhile, since there is no way for a subsystem to generate energy or particles (we only consider a single species of charge carrier here), the continuity of energy density and particle density respectively are:

$$0 = \frac{\partial u}{\partial t} + \nabla \cdot \vec{J}_U$$

$$0 = \frac{\partial n}{\partial t} + \nabla \cdot \vec{J}_N$$

Now, since the accumulation of entropy density can be expressed as the accumulation of energy and number density times inverse temperature and chemical potential over temperature respectively, we have:

$$\begin{aligned} \dot{s} &= \left( \frac{1}{T} \frac{\partial u}{\partial t} + \frac{-\mu_{ec}}{T} \frac{\partial n}{\partial t} \right) + \nabla \cdot \left( \frac{1}{T} \vec{J}_U + \frac{-\mu_{ec}}{T} \vec{J}_N \right) \\ &= \left( \nabla \frac{1}{T} \right) \cdot \vec{J}_U + \left( \nabla \frac{-\mu_{ec}}{T} \right) \cdot \vec{J}_N + \left( \frac{1}{T} \right) \left[ \frac{\partial u}{\partial t} + \nabla \cdot \vec{J}_U \right] + \left( \frac{-\mu_{ec}}{T} \right) \left[ \frac{\partial n}{\partial t} + \nabla \cdot \vec{J}_N \right] \\ &= \left( \nabla \frac{1}{T} \right) \cdot \vec{J}_U + \left( \nabla \frac{-\mu_{ec}}{T} \right) \cdot \vec{J}_N \end{aligned} \tag{B.1}$$

The rearrangement of energy density and particle density within regions of constant temperature and chemical potential do not constitute entropy generation, only

entropy accumulation, which is not entropically preferred for the system as a whole. Rather, the flow of a Joule of energy from a region of high temperature where it contributes a small amount to the overall system entropy, to a region of low temperature where it contributes a large amount to the overall system entropy, is entropically preferred. Likewise, the flow of a particle from high chemical potential to low generates entropy for an isothermal system.

## B.2 The Onsager Reciprocity

For a particular class of systems, known as purely resistive or memoryless, it is possible to express the fluxes at any time solely in terms of the affinities at that time. If we further limit ourselves to small deviations from equilibrium, we can assume that this relationship is linear - that each of the fluxes may be expressed as a sum of the various affinities, each with some constant linear coefficient. If we enumerate the fluxes by subscript  $p$  and the affinities by subscript  $q$ , we may express the assumption of a linear, purely resistive system as follows:

$$J_p = \sum_q L_{qp} \mathcal{F}_q$$

where  $\mathcal{F}$  is an affinity.

An exceedingly important result from the study of thermodynamic fluctuations is the Onsager Reciprocity. Onsager assumed that the evolution of a thermodynamic system in the presence of a random thermodynamic fluctuation followed the same laws as the system in the presence of an externally-imposed macroscopic deviation. That is, a single Green function describes both processes. To the system, there is no difference between a random fluctuation and a human-induced fluctuation.

After making this key assumption, we then consider the reaction of a system to a fluctuation in a particular extrinsic quantity, such as the energy density. We note that such a fluctuation could induce a change in another extrinsic quantity, such as the particle number density, after a time  $\tau$ . Assuming the microscopic laws that govern

the Green function for the original fluctuation obey time-reversal symmetry, setting up the system to begin with the observed deviation in particle number density would cause the system to exhibit the initial fluctuation again a time  $\tau$  *later*. Since this is true for all values of  $\tau$ , it must also be true for the derivatives with respect to this delay, fluxes, which we assumed previously were easily written in terms of affinities. Keeping in mind that we have assumed deviations and fluctuations are conceptually identical, we arrive at the conclusion that time-reversal symmetry requires that the relationship between one flux's affinities and a second flux are identical to the relationship between the second flux's affinities and the first flux.

This result is known as the Onsager Reciprocity:

$$L_{qp}(\vec{B}) = L_{pq}(-\vec{B})$$

where the dependence on magnetic field expresses the need to reverse magnetic field to retain time-reversal symmetry.

### B.3 The Onsager Relations

We will now seek to derive the fundamental

In this section, we present the Onsager relations which form the basis for the self-consistent Onsager-solver, and which permit us to apply the Onsager reciprocity to thermoelectricity in Section B.4. We begin with:

$$ds = \frac{1}{T} du - \frac{\mu_{ec}}{T} dn$$

and

$$J_Q \equiv J_U - \mu_{ec} J_N$$

then apply the results of previous sections to arrive at [3]:

$$\begin{aligned}
J_N &= -\left(\frac{T^2\sigma}{q^2}\right)\frac{1}{T}\nabla\mu_{ec} + \left(\frac{T^2\sigma\alpha}{q}\right)\nabla\frac{1}{T} \\
J_Q &= -\left(\frac{T^2\sigma\alpha}{q}\right)\frac{1}{T}\nabla\mu_{ec} + (T^3\sigma\alpha^2 + T^2\kappa)\nabla\frac{1}{T}
\end{aligned}$$

or

$$\begin{aligned}
\vec{J} &= \sigma\vec{E} - \sigma\alpha\nabla T \\
\vec{Q} &= \alpha T\vec{J} - \kappa\nabla T
\end{aligned}$$

## B.4 The Kelvin Relations

Set up the boundary conditions on the above equations to solve for the heat flow discontinuity under isothermal conditions- the phenomenology of the Peltier effect. We find that the Peltier coefficient, as defined in Figure 1-2 obeys

$$\Pi = \alpha \cdot T$$

Recalling what the Onsager reciprocity tells us, namely that  $L_{ij}$  is symmetric under zero magnetic field and obeys  $L_{ij}(B) = L_{ji}(-B)$  under non-zero magnetic field, for the case of the continuous system above with fluxes of particles and energy, this amounts to a slight modification of the equation above. The correct statement is known as the second Kelvin relation:

$$\Pi(B) = \alpha(-B) \cdot T$$

We may use this result to explain why a B-field in one direction may be good for a refrigerator while the opposite direction may be good for a generator [?], assuming materials with a big linear field dependence of the Seebeck can be found [?].

We should also setup the boundary conditions to identify the Thomson coefficient

and thereby achieve the First Kelvin Relation.

# Appendix C

## Simulation Source Code

### C.1 Numerical Solver for Microscopic Transport Coefficients Outside the Boltzmann Limit (MAT- LAB)

```
function [Sigma Sigma_Drude D D_Drude S S_Drude] = Sigma_D_and_S(mratio, tau, n, T,

hbar = (6.626e-34)/(2*pi);           % Planck's constant
kB = 1.3806503e-23;                 % Boltzmann's constant
mass = mratio*9.11e-31;             % carrier mass
q = -1.602e-19;                     % carrier charge

delta_n = n*0.001;
n_table = [n-delta_n n n+delta_n];
delta_T = T*0.001;
T_table = [T-delta_T T T+delta_T];
A = zeros(length(n_table), length(T_table));

for i=1:length(n_table)
    for j=1:length(T_table)
```

```

        A(i,j) = calculate_A(mratio, n_table(i), T_table(j));
    end
end
%A

if ( abs( (A(2,2)-A(1,2)) - (A(3,2)-A(2,2)) ) > (A(2,2)*1e-3) )
    delta_n_too_big = 1
end
dAdn = (A(3,2) - A(1,2))/(2*delta_n);
if ( abs( (A(2,2)-A(2,1)) - (A(2,3)-A(2,2)) ) > (A(2,2)*1e-3) )
    delta_T_too_big = 1
end
dAdT = (A(2,3) - A(2,1))/(2*delta_T);

%Integral for Conductivity
if (verbose == 1)
    Sigma = calculate_conductivity(mratio, tau, A(2,2), T)
    Sigma_Drude = q*n*(q*tau/mass)
else
    Sigma = calculate_conductivity(mratio, tau, A(2,2), T);
    Sigma_Drude = q*n*(q*tau/mass);
end

%Integral for Diffusivity
if (verbose == 1)
    D = calculate_diffusivity(mratio, tau, A(2,2), T, dAdn)
    D_Drude = (kB*T/q)*(q*tau/mass)
else
    D = calculate_diffusivity(mratio, tau, A(2,2), T, dAdn);
    D_Drude = (kB*T/q)*(q*tau/mass);
end

```



```

end

%Integral for Soret
if (verbose == 1)
    S = calculate_soret(mratio, tau, A(2,2), T, dAdT)
    S_Drude = n*(kB/q)*(q*tau/mass)
else
    S = calculate_soret(mratio, tau, A(2,2), T, dAdT);
    S_Drude = n*(kB/q)*(q*tau/mass);
end

extra_plot = false;
if (extra_plot)
    plot_num = 30;
    density = zeros(plot_num);
    A_of_density = zeros(plot_num);
    for b=1:plot_num
        density(b) = n*(1e-2)*(10)^((b-1)/10);
        A_of_density(b) = calculate_A(mratio, density(b), T);
    end
    figure;
    semilogx(density, (-kB*T*log(A_of_density))/abs(q), 'b+', 'MarkerSize', 5);
    title('Inverse Fermi Integral');
    xlabel('Density');
    ylabel('(E_F - E_C) in eV');
    set(gca, 'FontSize', 20);
    set(gca, 'LineWidth', 3);
end

end

```

```

% A is defined as exp( (E_C-E_F)/(kB*T) )
function [A] = calculate_A(mratio, n, T)

% CONSTANTS
hbar = (6.626e-34)/(2*pi);      % Planck's constant
kB = 1.3806503e-23;           % Boltzmann's constant
mass = mratio*9.11e-31;       % carrier mass
q = -1.602e-19;               % carrier charge

% INITIAL CONDITION
N_C = 2 * (mass*kB*T/(2*pi*hbar^2))^(3/2); % effective DOS at the band-edge
A_guess = N_C/n;              % starting guess for A
A = -1e6;                      % impossible value for A to be

% ITERATIVE CALCULATION PARAMETERS
k_thermal = (mass*kB*T / (hbar^2))^(1/2); % thermal wavevector
iter_limit = 2000;            % max # of iterations

for i=1:iter_limit
    k = linspace(0,(k_thermal*10),(k_thermal/1e3));
    integrand = (1/(pi^2)) * (k.^2) .* ( A_guess*exp((k.^2)/(2*k_thermal^2)) + 1 );
    n_of_A_guess = trapz(k, integrand);

    if (abs(n_of_A_guess - n) < n*1e-5)
        A = A_guess;
        break;
    else if (n_of_A_guess < n)
        A_guess = A_guess*(1 - (abs(n-n_of_A_guess)/n));
    end
end

```

```

        end
    end
end

if (A == -1e6)
    calculate_A_failed = 1
    A = A_guess
    proper_n_of_A = n_of_A_guess
    n_desired = n
end

end

function [Sigma] = calculate_conductivity(mratio, tau, A, T)

% CONSTANTS
hbar = (6.626e-34)/(2*pi);      % Planck's constant
kB = 1.3806503e-23;           % Boltzmann's constant
mass = mratio*9.11e-31;       % carrier mass
q = -1.602e-19;               % carrier charge

E_max = 1.1*max(0, -kB*T*log(A)) + 20*kB*T;
k_max = sqrt(2*mass*E_max)/(hbar);
if (k_max > (2*pi/(2.4e-10)))
    parabolicity_infeasible = 1
end

k_thermal = (mass*kB*T / (hbar^2))^(1/2); % thermal wavevector

kz = linspace(-k_max, k_max, 4001);
kperp = linspace(0, k_max, 2001);

```

```

f_0 = zeros(length(kz), length(kperp));
for i=1:length(kz)
    for j=1:length(kperp)
        f_0(i,j) = (1/(4*pi^3))*(1/( 1 + A*exp(((kz(i))^2 + (kperp(j))^2)/(2*k_ther
    end
end

%figure;
%subplot(2,1,1); plot(kperp, (4*pi^3)*f_0(501,:), 'LineWidth', 3);
%subplot(2,1,2); plot(kz, (4*pi^3)*f_0(:,1), 'LineWidth', 3);

%THIS DOESN'T WORK BECAUSE
%    tau(k) depends on kz and so must be left outside d/dkz[.]
%
%tau = zeros(length(kz), length(kperp));
%for i=1:length(kz)
%    for j=1:length(kperp)
%        tau(i,j) = tau_th*((kz(i))^2 + (kperp(j))^2)/(2*k_thermal^2) )^(p/2);
%    end
%end

%for i=1:length(kz)
%    for j=1:length(kperp)
%        kperp_matrix(i,j) = kperp(j);
%    end
%end

%integrand_2d = f_0 .* (tau*(-(q^2)/mass)) * ((2*pi) * kperp_matrix);

f_of_kz = zeros(size(kz));

```

```

for i=1:length(kz)
    f_of_kz(i) = trapz( kperp, (f_0(i,:) .* kperp) * (2*pi) * (tau*(-(q^2)/mass)) )
end

kzmid = conv(kz,[0.5 0.5]); kzmid = kzmid(2:end-1);
integrand = ( diff(f_of_kz)./diff(kz) ) .* kzmid;

%figure;
%plot(kzmid, integrand, 'r-', 'LineWidth', 2);

Sigma = trapz(kzmid, integrand);

end

function [D] = calculate_diffusivity(mratio, tau, A, T, dAdn)

% CONSTANTS
hbar = (6.626e-34)/(2*pi);      % Planck's constant
kB = 1.3806503e-23;           % Boltzmann's constant
mass = mratio*9.11e-31;       % carrier mass
q = -1.602e-19;               % carrier charge

E_max = 1.1*max(0, -kB*T*log(A)) + 20*kB*T;
k_max = sqrt(2*mass*E_max)/(hbar);
if (k_max > (2*pi/(2.4e-10)))
    parabolicity_infeasible = 1
end

k_thermal = (mass*kB*T / (hbar^2))^(1/2); % thermal wavevector

kz = linspace(-k_max, k_max, 4001);

```

```

kperp = linspace(0, k_max, 2001);

df0dn = zeros(length(kz), length(kperp));
for i=1:length(kz)
    for j=1:length(kperp)
        df0dn(i,j) = (-dAdn/(4*pi^3))*(exp(((kz(i))^2 + (kperp(j))^2)/(2*k_thermal'
    end
end

f_of_kz = zeros(size(kz));
for i=1:length(kz)
    f_of_kz(i) = trapz( kperp, (df0dn(i,:) .* kperp) * (2*pi) * (tau*(hbar/mass)^2)
end

integrand = f_of_kz .* (kz.^2);

%figure;
%plot(kz, integrand, 'r-', 'LineWidth', 2);

D = trapz(kz, integrand);

end

function [S] = calculate_soret(mratio, tau, A, T, dAdT)

% CONSTANTS
hbar = (6.626e-34)/(2*pi);           % Planck's constant
kB = 1.3806503e-23;                 % Boltzmann's constant
mass = mratio*9.11e-31;             % carrier mass
q = -1.602e-19;                     % carrier charge

```

```

E_max = 1.1*max(0, -kB*T*log(A)) + 20*kB*T;
k_max = sqrt(2*mass*E_max)/(hbar);
if (k_max > (2*pi/(2.4e-10)))
    parabolicity_infeasible = 1
end
k_thermal = (mass*kB*T / (hbar^2))^(1/2); % thermal wavevector

kz = linspace(-k_max, k_max, 4001);
kperp = linspace(0, k_max, 2001);

df0dT = zeros(length(kz), length(kperp));
for i=1:length(kz)
    for j=1:length(kperp)
        df0dT(i,j) = (-1/(4*pi^3))*(1/( 1 + A*exp(((kz(i))^2 + (kperp(j))^2)/(2*k_t
        df0dT(i,j) = df0dT(i,j)*(exp(((kz(i))^2 + (kperp(j))^2)/(2*k_thermal^2)))*
    end
end

f_of_kz = zeros(size(kz));
for i=1:length(kz)
    f_of_kz(i) = trapz( kperp, (df0dT(i,:) .* kperp) * (2*pi) * (tau*(hbar/mass)^2)
end

integrand = f_of_kz .* (kz.^2);

%figure;
%plot(kz, integrand, 'r-', 'LineWidth', 2);

S = trapz(kz, integrand);

```

```

end

function [] = coeff_plots()

T = 300;
tau = 3e-13;
mratio = 0.063;

num_densities = 31;
min_density = 1e22;
max_density = 1e25;
for i=1:num_densities
    density(i) = min_density * (max_density/min_density)^((i-1)/(num_densities-1));
end

for i=1:num_densities
    [out1 out2 out3 out4 out5 out6] = Sigma_D_and_S(mratio, tau, density(i), T, 0);
    sigma_real(i) = out1;
    sigma_drude(i) = out2;
    D_real(i) = out3;
    D_drude(i) = out4;
    S_real(i) = out5;
    S_drude(i) = out6;
end

figure;
loglog(density*1e-6, sigma_real/100, 'b-', 'MarkerSize', 10, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('Conductivity (A V-1 cm-1)', 'FontSize', 20);

```



```

set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

figure;
semilogx(density*1e-6, sigma_real./sigma_drude, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('(Real Conductivity / Boltzmann-Expression Conductivity) Ratio', 'FontSize',
set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

figure;
loglog(density*1e-6, D_real*1e4, 'b-', 'MarkerSize', 10, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('Diffusivity (cm2 s-1)', 'FontSize', 20);
set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

figure;
semilogx(density*1e-6, D_real./D_drude, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('(Real Diffusivity / Boltzmann-Expression Diffusivity) Ratio', 'FontSize', 2
set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

figure;
loglog(density*1e-6, S_real/100, 'b-', 'MarkerSize', 10, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('Soret (K-1 cm-1 s-1)', 'FontSize', 20);
set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

```

```

figure;
semilogx(density*1e-6, S_real./S_drude, 'LineWidth', 3);
xlabel('Carrier Density (cm-3)', 'FontSize', 20);
ylabel('(Real Soret / Boltzmann-Expression Soret) Ratio', 'FontSize', 20);
set(gca,'LineWidth', 2);
set(gca,'FontSize', 20);

end

```

## C.2 Seebeck Generalized Drift-Diffusion Solver (MATLAB)

```

function [field_seebeck diff_end_seebeck conductance_doping] = micro_seebeck_graded

% Program to self-consistently find the density profile of a TE element

%FUNDAMENTAL CONSTANTS
% all units MKS
q = -1.602e-19; %negative carriers
kB = 1.3806503e-23; %boltzmann's constant
eps0 = 1*(8.85418782e-12); %permittivity of free space

%MATERIAL PARAMETERS
eff_mass = 0.063*9.11e-31;
tau = 3e-13; %when tau = 1ps, then mobility ~ = 3 m2/Vs for m*=0.06
gamma_ratio = gamma((5+p)/2) / gamma(5/2);
temp_diff_coeff = gamma_ratio*(-kB*tau/eff_mass)*(1+(p/2)); % locally, J_T=(temp_di

```

```

diff_coeff = gamma_ratio*(-kB*tau/eff_mass); % locally, J_n=(diff_coeff)*T*gradn
mobility = gamma_ratio*(q/eff_mass)*tau;
% NOTE: TO CHANGE DIELECTRIC CONSTANT, WE MUST CHANGE poisson.m AND
% self_consistent_function.m as well
epsr = (3.255^2); %(GaAs epsr = n^2)
eps = epsr*eps0;

```

```

%INPUT PARAMETER INITIALIZATIONS

```

```

Th = T_in+0.5; % temp of hot side
Tc = T_in-0.5; % temp of cold side
T_avg = (Th+Tc)/2;
total_size = 1e-6;
num_grid_points = 401;
dx = total_size/(num_grid_points-1);
%dx = 0.00005;
%num_grid_points = (total_size / dx) + 1;
x = [0:dx:total_size];
xmid = conv(x,[.5 .5]); xmid = xmid(2:end-1);
T = linspace(Th, Tc, num_grid_points);
Tmid = conv(T,[.5 .5]); Tmid = Tmid(2:end-1);

```

```

%CREATE DOPANT PROFILE:

```

```

%{
n_dopant = linspace(0, 0, size(x,2));
%this is a method to produce a dopant profile with boundaries as passed and
%a center as passed. there should be a SMOOTH transition.
edge_length = round(num_grid_points/10);
for c=1:edge_length
    n_dopant(c) = n_edge;

```

```

    n_dopant(end-(c-1)) = n_edge;
end
for c=(edge_length+1):2*edge_length
    n_dopant(c) = n_edge + (n_bulk-n_edge)*((c-edge_length)/edge_length);
    n_dopant(end-(c-1)) = n_edge + (n_bulk-n_edge)*((c-edge_length)/edge_length);
end
for c=(2*edge_length):(num_grid_points-(2*edge_length))
    n_dopant(c) = n_bulk;
end
n_avg = sum(n_dopant)/num_grid_points;
n = n_dopant;
nmid = conv(n,[.5 .5]); nmid = nmid(2:end-1);
%}
n_dopant = linspace(0, 0, size(x,2));
%this is a method to produce a dopant profile with boundaries as passed and
%a center as passed. there should be a SHARP transition.
edge_length = round(num_grid_points/10);
for c=1:edge_length
    n_dopant(c) = n_edge;
    n_dopant(end-(c-1)) = n_edge;
end
for c=edge_length:(num_grid_points-edge_length)
    n_dopant(c) = n_bulk;
end
n_avg = sum(n_dopant)/num_grid_points;
n = n_dopant;
nmid = conv(n,[.5 .5]); nmid = nmid(2:end-1);

%{
%this is a method to create graded n+ boundaries

```

```

n_avg = 1e20;
n_dopant = linspace(n_avg, n_avg, size(x,2));
edge=20;
dope_slope=0.5/edge;
for c=1:edge
    n_dopant(c) = n_dopant(c) + n_avg*dope_slope*(c-edge);
    n_dopant(end-(c-1)) = n_dopant(end-(c-1)) + n_avg*dope_slope*(c-edge);
end
for c=(edge+1):(num_grid_points-edge)
    n_dopant(c) = ( n_avg*num_grid_points - sum(n_dopant(1:edge)+n_dopant(end-edge)
end
%}

%{
%this is a method to create n+ boundaries with constant dopant concentration
emitter_length = round(num_grid_points/10);
for c=1:emitter_length
    n_dopant(c) = n_avg*5;
end
for c=emitter_length:num_grid_points
    n_dopant(c) = (n_avg*num_grid_points - sum(n_dopant(1:emitter_length))) / (num_
end
n = n_dopant;
nmid = conv(n, [.5 .5]); nmid = nmid(2:end-1);
%}

%FLOW INITIALIZATIONS
J_T = zeros(1,(num_grid_points-1));
J_n = zeros(1,(num_grid_points-1));
J_E = zeros(1,(num_grid_points-1));

```

```

%GENERAL SOLUTION INITIALIZATIONS

numits = 12; % max number of iterations for self-consistent loop
branch = 4; % branching factor for newton's method
residue = zeros(num_grid_points,1);
jacobian = zeros(num_grid_points, num_grid_points);
modsq = sum(residue.*residue); % modulus squared of residue vector

% BRANCHING TEMP VARIABLE INITIALIZATIONS
delta_n = linspace(0, 0, num_grid_points); % n_new = n_old + delta_n
new_n = linspace(0, 0, num_grid_points);
new_nmid = conv(new_n,[.5 .5]); new_nmid = new_nmid(2:end-1);
new_T = linspace(0, 0, num_grid_points);
new_Tmid = conv(new_T,[.5 .5]); new_Tmid = new_Tmid(2:end-1);
new_residue = zeros(num_grid_points,1);
new_J_n = zeros(1,(num_grid_points-1));
new_J_T = zeros(1,(num_grid_points-1));
new_J_E = zeros(1,(num_grid_points-1));
new_modsq = zeros(1,branch);

% RECORD-KEEPING VARIABLES
damping_record = zeros(1,numits);
modsq_record = zeros(1,numits);
all_modsq_record = zeros(branch,numits);
n_record = zeros(numits+1,num_grid_points);

%J_T = n_avg*temp_diff_coeff.*(diff(T)./diff(x));

%initial guess at J_T -> J_n -> n
J_T = temp_diff_coeff*nmid.*(diff(T)./diff(x));

```

```

J_n = -J_T;
n(1)=0;
for i=2:num_grid_points
    n(i) = n(i-1) + (dx * (J_n(i-1) / (Tmid(i-1)*diff_coeff)));
end
n = n + (sum(n_dopant-n)/num_grid_points);
n_record(1,:) = n;
%n = n_outside;

success = 0;
for j=1:numits
    % calculate residue
    nmid = conv(n,[.5 .5]); nmid = nmid(2:end-1);
    Tmid = conv(T,[.5 .5]); Tmid = Tmid(2:end-1);
    for k=1:(num_grid_points-1)
        J_E(k) = (q*mobility*dx/eps)*sum(nmid(1:k)-n_dopant(1:k))*nmid(k); %E-field
        J_T(k) = (temp_diff_coeff*nmid(k))*(T(k+1)-T(k))/(x(k+1)-x(k));
        J_n(k) = (diff_coeff*Tmid(k))*(n(k+1)-n(k))/(x(k+1)-x(k));
    end
    residue(1:end-1) = J_T + J_n + J_E;
    residue(end) = (1e0)*sum(n-n_dopant);
    if (mod(j,3)==0)
        modsq = sum(residue.*residue); %remove semicolon to print modsq every 3rd i
    else
        modsq = sum(residue.*residue);
    end
    modsq_record(j) = modsq;
    if (j == 1)
        init_modsq = modsq;
    end
end

```

```

if (modsq < ((1e-20)*init_modsq))
    success = 1;
    break;
end

% calculate jacobian
jacobian = zeros(num_grid_points, num_grid_points);
for a=1:(num_grid_points-1)
    for b=1:num_grid_points
        % fill in J_E terms for jacobian
        if (b <= a)
            jacobian(a,b) = jacobian(a,b) + nmid(a)*(q*mobility*dx/eps);
        end
        if ((b == a) || (b == (a+1)))
            jacobian(a,b) = jacobian(a,b) + (0.5)*sum(n(1:a)-n_dopant(1:a))*(q*
        end
        % fill in J_n terms for jacobian
        if (b == a)
            jacobian(a,b) = jacobian(a,b) - (diff_coeff/dx)*Tmid(a);
        end
        if (b == (a+1))
            jacobian(a,b) = jacobian(a,b) + (diff_coeff/dx)*Tmid(a);
        end
        % fill in J_T terms for jacobian
        if (b == a)
            jacobian(a,b) = jacobian(a,b) + (temp_diff_coeff/2)*((T(a+1)-T(a)),
        end
        if (b == (a+1))
            jacobian(a,b) = jacobian(a,b) + (temp_diff_coeff/2)*((T(a+1)-T(a)),
        end
    end
end

```



```

        end
    end
    jacobian(end,:) = ones(1,num_grid_points);

    % take step
    new_modsq = ones(1,branch).*(modsq);
    for c=1:branch
        % set up damping value for this branch
        damping = (0.9)^(c-1);

        % calculate new residue for this branch
        delta_n = (-1)*damping*inv(jacobian)*residue;
        new_n = n + delta_n';
        new_T = T;
        new_nmid = conv(new_n,[.5 .5]); new_nmid = new_nmid(2:end-1);
        new_Tmid = conv(new_T,[.5 .5]); new_Tmid = new_Tmid(2:end-1);
        for k=1:(num_grid_points-1)
            new_J_E(k) = (q*mobility*dx/eps)*new_nmid(k)*sum(new_nmid(1:k)-n_dopant);
            new_J_T(k) = (temp_diff_coeff*new_nmid(k))*(new_T(k+1)-new_T(k))/(x(k+1)-x(k));
            new_J_n(k) = (diff_coeff*new_Tmid(k))*(new_n(k+1)-new_n(k))/(x(k+1)-x(k));
        end
        new_residue(1:end-1) = new_J_T + new_J_n + new_J_E;
        new_residue(end) = (1e0)*sum(new_n-n_dopant);

        new_modsq(c) = sum(new_residue.*new_residue);
        all_modsq_record(c,j) = new_modsq(c);

        if (c == branch)
            [min_val,min_loc] = min(new_modsq);
            if (min_val >= modsq)

```

```

        damping_record(j) = (0.9)^(min_loc-1);
        error = 'no branch attempted could improve |residue|^2'
        break;
    else
        damping = (0.9)^(min_loc-1);
        damping_record(j) = damping;
        delta_n = (-1)*damping*inv(jacobian)*residue;
        n = n + delta_n';
    end
    n_record(j+1,:) = n;
end
end
end
% if ((min_val >= modsq) && (c == branch))
%     break;
% end
end

%post-processing for output
E_guess = linspace(0, 0, num_grid_points);
E_guess = poisson(n,n_dopant,q,x);
[field_seebeck_self_consistent ECminusEF_self_consistent ECminusEF_orig] = self_cor
if (success == 1)
    field_seebeck = (sum(-1*dx*E_guess) - field_seebeck_self_consistent)/(Th-Tc);
    DOSbandedgeT = (((pi^(3/2))*sqrt(2))^( -1))*((eff_mass*kB*T)/(((6.626e-34)/(2*pi
%figure; plot(DOSbandedgeT);
    ECminusEF = -kB*T.*log(n./DOSbandedgeT);
    %ECminusEF_orig = -kB*T.*log(n_dopant./DOSbandedgeT);
    %above NO LONGER USED because uses new T with n_dopant (before Poisson)
%{

```

```

if (toggle_plots == 1)
    figure; plot(x,ECminusEF/abs(q),'r','LineWidth',3); ylim([max(ECminusEF/abs
    hold on; plot(x,ECminusEF_self_consistent/abs(q),'b','LineWidth',3);
    hold on; plot(x,ECminusEF_orig/abs(q),'g','LineWidth',3);
    hold on; plot(x,0,'k:','LineWidth',3);
    set(gca,'LineWidth',3); set(gca,'FontSize',20); xlabel('position'); ylabel(
    legend('E_C After \nabla T is Applied', 'E_C Before \nabla T is Applied but
end
%}

diff_end_seebeck = ((ECminusEF(end)-ECminusEF_self_consistent(end))-(ECminusEF(
seebeck = diff_end_seebeck + field_seebeck;
peltier = seebeck * ((Tc+Th)/2);
conductance_doping = (sum(1./n)/num_grid_points)^(-1); %this quantity is the ef
power_factor = (seebeck)^2 * (conductance_doping * mobility * q);
end

if ((success == 1) && (toggle_plots == 1))
    EC = zeros(1,num_grid_points);
    EC(1) = 0;
    for i=2:num_grid_points
        EC(i) = EC(i-1) + abs(q)*(E_guess(i-1)+E_guess(i))*(0.5)*dx; %note that the
    end

%figure with EC and EF in subplots
figure;
subplot(2,1,1); plot(1e9*x,1000*EC/abs(q),'g','LineWidth',3); hold on;
xlim(1e9*[x(1) x(end)]);
xlabel('Position (nm)', 'FontSize', 20); ylabel('E_C and E_F (meV)', 'FontSize'
title('Conduction Band', 'FontSize', 20);
set(gca,'LineWidth',3); set(gca,'FontSize',20);

```

```

subplot(2,1,2); plot(1e9*x,1000*(EC-ECminusEF)/abs(q),'b','LineWidth',3);
xlim(1e9*[x(1) x(end)]);
xlabel('Position (nm)', 'FontSize', 20); ylabel('E_C and E_F (meV)', 'FontSize',
title('Fermi Level', 'FontSize', 20);
set(gca,'LineWidth',3); set(gca,'FontSize',20);

%{
%figure with EC and EF
figure; plot(x,EC/abs(q),'g','LineWidth',3);
hold on; plot(x,(EC-ECminusEF)/abs(q),'b','LineWidth',3);
ylim([(max(ECminusEF/abs(q))+max(abs(EC)/abs(q)))*(-1.1)-0.1 max(abs(EC)/abs(q)
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('E_C and E_F (electron-Volts)');
legend('E_C', 'E_F', 'Location', 'B');
%figure with just EF
figure; plot(x,(EC-ECminusEF)/abs(q),'b','LineWidth',3);
ylim([(min((EC-ECminusEF)/abs(q)))-0.002 max((EC-ECminusEF)/abs(q))+0.002]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('E_F (electron-Volts)');
%}

%{
% NEXT 4 PLOTS TRYING TO FIGURE OUT CENTER DIFFUSION CONTRIB TO grad(EF-EC)
%figure with density gradients
figure; plot(xmid,(n(2:end)-n(1:end-1))./((n(2:end)+n(1:end-1))/2),'b','LineWid
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('\Delta n / n (fraction)');
title('Fractional Density Change from Gridpoint to Gridpoint');
%figure with density gradients
figure; plot(xmid,log((n(2:end)-n(1:end-1))./((n(2:end)+n(1:end-1))/2)),'r','Li

```

```

set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('\Delta n / n (log of fraction)');
title('Log Fractional Density Change from Gridpoint to Gridpoint');
%figure with density gradients
figure; plot(xmid,n(2:end)./n(1:end-1),'b','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('\Delta n / n (fraction)');
title('Density Ratio from Gridpoint to Gridpoint');
%figure with density gradients
figure; plot(xmid,log(n(2:end)./n(1:end-1)),'r','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('\Delta n / n (log of fraction)');
title('Log Density Ratio from Gridpoint to Gridpoint');
%}

```

```

%figure with just n and n_D
figure;
plot(1e9*x,n_dopant*1e-6,'r--','LineWidth',5); hold on;
plot(1e9*x,n*1e-6,'b','LineWidth',4);
xlim(1e9*[x(1) x(end)]);
ylim([max(n_dopant*1e-6)*0.9997 max(n_dopant*1e-6)*1.0003]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
title(['Carrier Density Profile, T=300.5K']);
xlabel('Position (nm)'); ylabel('Number Density (cm-3)');
xlabelposition = 1e9*x(round(num_grid_points/20));
ylabelposition = (n_dopant(round(num_grid_points/20))*1e-6)+(1e-6)*(n_dopant(ro
text(xlabelposition, ylabelposition, 'n_D \newline \downarrow', 'FontSize', 30,
text(1e9*x(end-round(num_grid_points/20)), (1e-6)*n(end-round(num_grid_points/2
%legend('Dopant Density', 'Carrier Density', 'Location', 'B'));

```

```

%figure with delta n (and associated current)
%{
figure; subplot(2,1,1);
plot(x,n,'b','LineWidth',3);
hold on; plot(x,n_dopant,'k:','LineWidth',3);
ylim([max(n_dopant)*0.9997 max(n_dopant)*1.0003]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
title('Carrier Density and Associated Current for n-GaAs ( $n_{\text{dopant}} = 1e15 \text{ cm}^{-3}$ )');
xlabel('Position (m)'); ylabel('Number Density ( $\text{m}^{-3}$ )');
legend('Carrier Density', 'Dopant Density', 'Location', 'B');
subplot(2,1,2);
plot(xmid,J_n,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('Particle Flux ( $\# \cdot \text{m}^{-2} \text{ s}^{-1}$ )');
%}

%figure with E-field (and associated current)
%{
figure; subplot(2,1,1);
plot(x,E_guess,'g:','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
title('Electric Field and Associated Current for n-GaAs ( $n_{\text{dopant}} = 1e15 \text{ cm}^{-3}$ )');
xlabel('Position (m)'); ylabel('Electric Field (V/m)');
subplot(2,1,2);
plot(xmid,J_E,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('Particle Flux ( $\# \cdot \text{m}^{-2} \text{ s}^{-1}$ )');
%}

```

```

%figure with currents on same plot
figure;
plot(1e9*xmid,q*J_E*1e-6,'r-', 'LineWidth',2); hold on;
plot(1e9*xmid,q*J_n*1e-6,'g-', 'LineWidth',3); hold on;
plot(1e9*xmid,q*J_T*1e-6,'b-', 'LineWidth',4); hold on;
plot(1e9*xmid,q*(J_E + J_n + J_T)*1e-6,'k--', 'LineWidth',4);
set(gca, 'LineWidth',3); set(gca, 'FontSize',20);
flux_range3 = max(abs(q*J_T*1e-6));
xlim(1e9*[x(1) x(end)]);
ylim([-1.1*flux_range3 1.1*flux_range3]);
xlabel('Position (nm)'); ylabel('Current Density (A cm-2)');
title('Current Density By Source');
xlabelposE = 1e9*xmid(round(num_grid_points*0.6));
ylabelposE = q*1e-6*J_E(round(num_grid_points*0.6))*0.95;
text(xlabelposE, ylabelposE, '\uparrow J_E', 'VerticalAlignment', 'top', 'FontS
xlabelposn = 1e9*xmid(round(num_grid_points*0.08));
ylabelposn = q*1e-6*J_n(round(num_grid_points*0.08));
text(xlabelposn, ylabelposn, '\leftarrow J_{\nabla n}', 'FontSize', 30, 'Color',
xlabelposT = 1e9*xmid(round(num_grid_points*0.2));
ylabelposT = q*1e-6*J_T(round(num_grid_points*0.2));
text(xlabelposT, ylabelposT, '\downarrow J_{\nabla T}', 'FontSize', 30, 'Color',
xlabelpostot = 1e9*xmid(round(num_grid_points*0.94));
ylabelpostot = q*1e-6*(J_E(round(num_grid_points*0.94))+J_n(round(num_grid_poin
text(xlabelpostot, ylabelpostot, 'J_{tot} \uparrow', 'VerticalAlignment', 'top'

%legend('J_E', 'J_{\nabla n}', 'J_{\nabla T}', 'J_{total}', 'Location', 'B');

%figure with currents (in subplots)

```

```

%{
flux_range = max(abs(J_E)) - min(abs(J_E));
flux_range2 = max(abs(J_T));
figure; subplot(2,2,1);
plot(xmid,J_E,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
ylim([-1.1*flux_range2 0]);
xlabel('Position (m)'); ylabel('Particle Flux (# \cdot m^{-2} s^{-1})');
title('E Driven Current Density (J_E)');
subplot(2,2,2);
plot(xmid,J_n,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
ylim([-1.1*flux_range2 0]);
xlabel('Position (m)'); ylabel('Particle Flux (# \cdot m^{-2} s^{-1})');
title('\nabla Driven Current Density (J_{\nabla})');
subplot(2,2,3);
plot(xmid,J_T,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
ylim([0 1.1*flux_range2]);
xlabel('Position (m)'); ylabel('Particle Flux (# \cdot m^{-2} s^{-1})');
title('\nablaT Driven Current Density (J_{\nablaT})');
subplot(2,2,4);
plot(xmid,J_E + J_n + J_T,'m--','LineWidth',3);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
ylim([-0.55*flux_range2 0.55*flux_range2]);
xlabel('Position (m)'); ylabel('Particle Flux (# \cdot m^{-2} s^{-1})');
title('Net Current Density (J_E+J_{\nabla}+J_{\nablaT})');
%}

```

end



```

%{
if (toggle_plots == 1)
    %plot figures for output
    figure;

    subplot(3,2,1);
    plot(x,T,'r-','LineWidth',3); hold on;
    title('T(x)');
    set(gca,'FontSize',16);

    subplot(3,2,2);
    plot(xmid,J_T,'k','LineWidth',3);
    title('J_T(x)');
    set(gca,'FontSize',16);
    set(gca,'LineWidth',3);

    subplot(3,2,3);
    plot(x,n-n_dopant,'b','LineWidth',3); hold on;
    %plot(x,sum(n_dopant)/num_grid_points,'k:','LineWidth',2);
    %ylim([max([n_avg (n-n_dopant)])*(-0.1) max([n_avg (n-n_dopant)])*(1.1)]);
    title('\Delta n(x)');
    set(gca,'FontSize',16);

    subplot(3,2,4);
    plot(xmid,J_n,'k','LineWidth',3);
    title('J_n(x)');
    set(gca,'FontSize',16);
    set(gca,'LineWidth',3);

```

```

subplot(3,2,5);
plot(x,E_guess,'g:','LineWidth',3);
title('E(x)');
set(gca,'FontSize',16);

subplot(3,2,6);
plot(xmid,J_E,'k','LineWidth',3);
% hold on; plot(x,E_guess.*n*mobility,'r','LineWidth',3);
title('J_E(x)');
set(gca,'FontSize',16);
set(gca,'LineWidth',3);
end
%}

```

### C.3 Peltier Generalized Drift-Diffusion Solver (MATLAB)

```

function [n qJdotE xmid modsq_record] = peltier(n_L, n_C, n_R, T_in, qJ_applied, p,

% Program to self-consistently find the density profile of a semiconductor
% homojunction (unipolar) with a current flowing.

%FUNDAMENTAL CONSTANTS
% all units MKS
q = -1.602e-19; %negative carriers
kB = 1.3806503e-23; %boltzmann's constant
eps0 = 1*(8.85418782e-12); %permittivity of free space

%MATERIAL PARAMETERS

```

```

eff_mass = 0.063*9.11e-31;
tau = 3e-13; %when tau = 1ps, then mobility ~ = 3 m^2/Vs for m*=0.06
gamma_ratio = gamma((5+p)/2) / gamma(5/2);
temp_diff_coeff = gamma_ratio*(-kB*tau/eff_mass)*(1+(p/2)); % locally, J_T=(temp_diff
diff_coeff = gamma_ratio*(-kB*tau/eff_mass); % locally, J_n=(diff_coeff)*T*gradn
mobility = gamma_ratio*(q/eff_mass)*tau;
% NOTE: TO CHANGE DIELECTRIC CONSTANT, WE MUST CHANGE poisson.m AND
% self_consistent_function.m as well
%epsr = (3.255^2); %(GaAs epsr = n^2)
epsr = 1; %(GaAs epsr = n^2)
%epsr = 21;
eps = epsr*eps0;

%MESH GENERATION AND BOUNDARY CONDITIONS
Th = T_in; % temp of hot side
Tc = T_in; % temp of cold side
T_avg = (Th+Tc)/2;
total_size = 1e-6;
num_grid_points = 401;
dx = total_size/(num_grid_points-1);
%dx = 0.00005;
%num_grid_points = (total_size / dx) + 1;
x = [0:dx:total_size];
xmid = conv(x,[.5 .5]); xmid = xmid(2:end-1);
T = linspace(Th, Tc, num_grid_points);
Tmid = conv(T,[.5 .5]); Tmid = Tmid(2:end-1);

%CREATE DOPANT PROFILE:
n_dopant = linspace(0, 0, size(x,2));
%this is a method to produce a dopant profile

```

```

%[...]n_L...|...n_C...|...n_R...]
% there should be SHARP transitions.
boundary_positionA = round(num_grid_points/3);
boundary_positionB = round(2*num_grid_points/3);
for c=1:boundary_positionA
    n_dopant(c) = n_L;
end
for c=(boundary_positionA+1):boundary_positionB
    n_dopant(c) = n_C;
end
for c=(boundary_positionB+1):num_grid_points
    n_dopant(c) = n_R;
end
n_avg = sum(n_dopant)/num_grid_points;
n = n_dopant;
conductivity = q*(1/(mean(1./n)))*mobility;
E_applied = qJ_applied / conductivity;
n(round(num_grid_points/2)) = n(round(num_grid_points/2))*1.1;
nmid = conv(n,[.5 .5]); nmid = nmid(2:end-1);
n_dopant_mid = conv(n_dopant, [.5 .5]); n_dopant_mid = n_dopant_mid(2:end-1);
%{
if (toggle_plots == 1)
    figure; plot(x,n_dopant);
    ylim([min(n_dopant)-(0.1*(max(n_dopant)-min(n_dopant))) max(n_dopant)+(0.1*(max
end
%}

%FLOW INITIALIZATIONS
J_T = zeros(1,(num_grid_points-1));
J_n = zeros(1,(num_grid_points-1));

```

```

J_E = zeros(1,(num_grid_points-1));

%GENERAL SOLUTION INITIALIZATIONS
numits = 12; % max number of iterations for self-consistent loop
branch = 4; % branching factor for newton's method
residue = zeros(num_grid_points,1);
jacobian = zeros(num_grid_points, num_grid_points);
modsq = sum(residue.*residue); % modulus squared of residue vector

% BRANCHING TEMP VARIABLE INITIALIZATIONS
delta_n = linspace(0, 0, num_grid_points); % n_new = n_old + delta_n
new_n = linspace(0, 0, num_grid_points);
new_nmid = conv(new_n,[.5 .5]); new_nmid = new_nmid(2:end-1);
new_T = linspace(0, 0, num_grid_points);
new_Tmid = conv(new_T,[.5 .5]); new_Tmid = new_Tmid(2:end-1);
new_residue = zeros(num_grid_points,1);
new_J_n = zeros(1,(num_grid_points-1));
new_J_T = zeros(1,(num_grid_points-1));
new_J_E = zeros(1,(num_grid_points-1));
new_modsq = zeros(1,branch);

% RECORD-KEEPING VARIABLES
damping_record = zeros(1,numits);
modsq_record = zeros(1,numits);
all_modsq_record = zeros(branch,numits);
n_record = zeros(numits+1,num_grid_points);

%initial guess at J_applied -> J_n,J_E -> n
%{

```

```

J_T = temp_diff_coeff*nmid.*(diff(T)./diff(x)); %J_T is irrelevant
J_n = -J_applied;
n(1)=0;
for i=2:num_grid_points
    n(i) = n(i-1) + (dx * (J_n(i-1) / (Tmid(i-1)*diff_coeff)));
end
n = n + (sum(n_dopant-n)/num_grid_points);
n = n_dopant;
n_record(1,:) = n;
%n = n_outside;
%}

success = 0;
for j=1:numits
    % calculate residue
    nmid = conv(n,[.5 .5]); nmid = nmid(2:end-1);
    Tmid = conv(T,[.5 .5]); Tmid = Tmid(2:end-1);
    for k=1:(num_grid_points-1)
        J_E(k) = ((q*dx/eps)*sum(n(1:k)-n_dopant(1:k)) + E_applied)*mobility*nmid(k)
        J_T(k) = (temp_diff_coeff*nmid(k))*(T(k+1)-T(k))/(x(k+1)-x(k));
        J_n(k) = (diff_coeff*Tmid(k))*(n(k+1)-n(k))/(x(k+1)-x(k));
    end
    residue(1:end-1) = J_T + J_n + J_E - (qJ_applied/q);
    residue(end) = (1e0)*sum(n-n_dopant);
    %residue(end)
    if (mod(j,3)==0)
        modsq = sum(residue.*residue); %remove semicolon to print modsq every 3rd :
    else
        modsq = sum(residue.*residue);
    end
end

```

```

modsq_record(j) = modsq;
if (j == 1)
    init_modsq = modsq;
end
if (modsq < ((1e-27)*init_modsq))
    success = 1
    %figure; plot(J_E,'b', 'LineWidth', 3);
    %hold on; plot(J_n, 'r', 'LineWidth', 3);
    %hold on; plot(J_E+J_n, 'k--', 'LineWidth', 3);
    break;
end

% calculate jacobian
jacobian = zeros(num_grid_points, num_grid_points);
for a=1:(num_grid_points-1)
    for b=1:num_grid_points
        % fill in J_E terms for jacobian
        if (b <= a)
            jacobian(a,b) = jacobian(a,b) + (q*dx/eps)*mobility*nmid(a);
        end
        if ((b == a) || (b == (a+1)))
            jacobian(a,b) = jacobian(a,b) + ((q*dx/eps)*sum(n(1:a)-n_dopant(1:a)));
        end
        % fill in J_n terms for jacobian
        if (b == a)
            jacobian(a,b) = jacobian(a,b) - (diff_coeff/dx)*Tmid(a);
        end
        if (b == (a+1))
            jacobian(a,b) = jacobian(a,b) + (diff_coeff/dx)*Tmid(a);
        end
    end
end

```

```

        % fill in J_T terms for jacobian
        if (b == a)
            jacobian(a,b) = jacobian(a,b) + (temp_diff_coeff/2)*((T(a+1)-T(a))/
        end
        if (b == (a+1))
            jacobian(a,b) = jacobian(a,b) + (temp_diff_coeff/2)*((T(a+1)-T(a))/
        end
    end
end
jacobian(end,:) = ones(1,num_grid_points);

% take step
%PROBLEM IF IT GETS TO HERE EVEN AFTER SUCCEEDING.
%   if (success == 1)
%       problem = 1
%   end
new_modsq = ones(1,branch).*(modsq);
for c=1:branch
    % set up damping value for this branch
    damping = (0.6)^(c-1);

    % calculate new residue for this branch
    delta_n = (-1)*damping*inv(jacobian)*residue;
    new_n = n + delta_n';
    new_T = T;
    new_nmid = conv(new_n,[.5 .5]); new_nmid = new_nmid(2:end-1);
    new_Tmid = conv(new_T,[.5 .5]); new_Tmid = new_Tmid(2:end-1);
    for k=1:(num_grid_points-1)
        new_J_E(k) = ((q*dx/eps)*sum(new_n(1:k)-n_dopant(1:k)) + E_applied)*mol
        new_J_T(k) = (temp_diff_coeff*new_nmid(k))*(new_T(k+1)-new_T(k))/(x(k+

```



```

        new_J_n(k) = (diff_coeff*new_Tmid(k))*(new_n(k+1)-new_n(k))/(x(k+1)-x(k))
    end
    new_residue(1:end-1) = new_J_T + new_J_n + new_J_E - (qJ_applied/q);
    new_residue(end) = (1e0)*sum(new_n-n_dopant);

    new_modsq(c) = sum(new_residue.*new_residue);
    all_modsq_record(c,j) = new_modsq(c);

    if (c == branch)
        [min_val,min_loc] = min(new_modsq);
        if (min_val >= modsq)
            damping_record(j) = (0.9)^(min_loc-1);
            error = 'no branch attempted could improve |residue|^2'
            break;
        else
            damping = (0.9)^(min_loc-1);
            damping_record(j) = damping;
            delta_n = (-1)*damping*inv(jacobian)*residue;
            n = n + delta_n';
        end
        n_record(j+1,:) = n;
    end
end

% if ((min_val >= modsq) && (c == branch))
%     break;
% end
end

%post-processing for output

```

```

if (success == 1)
    %calculate electric field based on output density profile
    Efield = linspace(0, 0, num_grid_points);
    Efield = poisson_mid(n,n_dopant,eps,q,x) + E_applied;
%   figure; plot(xmid,Efield); hold on; plot(xmid,E_applied);

    EC = zeros(1,num_grid_points);
    EC(1) = 0;
    for i=1:length(Efield)
        EC(i+1) = EC(i) + abs(q)*(Efield(i))*dx; %note that the sign convention is
    end

    DOSbandedgeT = (((pi^(3/2))*sqrt(2))^(1))*((eff_mass*kB*T)/(((6.626e-34)/(2*pi
%   plot(DOSbandedgeT);
%   DOSbandedgeT = DOSbandedgeT*2;
    ECminusEF = -kB*T.*log(n./DOSbandedgeT);
    grad_mu = diff(-1*ECminusEF)./diff(x);

    J_solved_for = J_E + J_n + J_T;
    qJdotE = q*J_solved_for .* Efield; %Total Heat
    qJ_applied_dotE = qJ_applied * Efield;
    qJdot_grad_mu_over_q = q*J_solved_for .* grad_mu / q; %Peltier Heat
end

if ((success == 1) && (toggle_plots == 1))
    %figure with J dot E
    %PELTIER PLOT
    %{
    figure; %DEBUGGING
    subplot(3,1,1); plot(xmid,J_E); hold on; plot(xmid, J_applied);

```

```

subplot(3,1,2); plot(xmid,J_n); hold on; plot(xmid, J_applied);
subplot(3,1,3); plot(xmid,J_T); hold on; plot(xmid, J_applied);
%}

figure;
%plot(xmid, qJdotE, 'b', 'LineWidth', 4); hold on;
%plot(xmid, qJdot_grad_mu_over_q, 'b', 'LineWidth', 4); hold on;
plot(xmid, qJdotE - qJdot_grad_mu_over_q, 'b', 'LineWidth', 4); hold on;
plot(xmid, 0.*xmid, 'k--', 'LineWidth', 2);
%plot(1e9*xmid, qJ_applied_dotE, 'r--', 'LineWidth', 2); hold on;
xlim([x(1) x(end)]);
xlabel('Position (m)', 'FontSize', 20); ylabel('J dot E (W m-3)', 'FontS
title('-- Heat Density Production from -- Current', 'FontSize', 25);
set(gca,'LineWidth',3); set(gca,'Fontsize',20);

%figure with EC and EF in subplots
%PELTIER PLOT
%{
figure;
subplot(2,1,1); plot(1e9*x,1000*EC/abs(q),'g','LineWidth',3); hold on;
xlim(1e9*[x(1) x(end)]);
xlabel('Position (nm)', 'FontSize', 20); ylabel('E_C (meV)', 'FontSize', 20);
title('Conduction Band', 'FontSize', 20);
set(gca,'LineWidth',3); set(gca,'Fontsize',20);
subplot(2,1,2); plot(1e9*x,1000*(EC-ECminusEF)/abs(q),'b','LineWidth',3);
xlim(1e9*[x(1) x(end)]);
Escale = min(abs(max(ECminusEF)-min(ECminusEF))*1000/abs(q), (1000*kB*T_in/abs(
ylim([(1000*mean(EC-ECminusEF)/abs(q))-Escale (1000*mean(EC-ECminusEF)/abs(q))+
xlabel('Position (nm)', 'FontSize', 20); ylabel('E_F (meV)', 'FontSize', 20);

```

```

title('Fermi Level', 'FontSize', 20);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
%}

%figure with EC and EF
%PELTIER PLOT
%{
figure; plot(x,EC/abs(q),'g','LineWidth',3);
hold on; plot(x,(EC-ECminusEF)/abs(q),'b','LineWidth',3);
ylim([(max(ECminusEF/abs(q))+max(abs(EC)/abs(q)))*(-1.1)-0.1 max(abs(EC)/abs(q))]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('E_C and E_F (electron-Volts)');
legend('E_C', 'E_F', 'Location', 'B');
%figure with just EF
figure; plot(x,(EC-ECminusEF)/abs(q),'b','LineWidth',3);
ylim([(min((EC-ECminusEF)/abs(q)))-0.002 max((EC-ECminusEF)/abs(q))+0.002]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
xlabel('Position (m)'); ylabel('E_F (electron-Volts)');
%}

%figure with just n and n_D
%PELTIER PLOT
%{
figure;
plot(1e9*x,n_dopant*1e-6,'r--','LineWidth',4); hold on;
plot(1e9*x,n*1e-6,'b','LineWidth',4);
xlim(1e9*[x(1) x(end)]);
ylim([min(n_dopant*1e-6)*0.9 max(n_dopant*1e-6)*1.1]);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
title(['Carrier Density Profile, T=300.5K']);

```

```

xlabel('Position (nm)'); ylabel('Number Density (cm-3)');
xlabelposition = 1e9*x(round(num_grid_points/20));
ylabelposition = (n_dopant(round(num_grid_points/20))*1e-6)+(1e-6)*(n_dopant(rc
text(xlabelposition, ylabelposition, 'n_D \newline \downarrow', 'FontSize', 30,
text(1e9*x(end-round(num_grid_points/20)), (1e-6)*n(end-round(num_grid_points/2
%legend('Dopant Density', 'Carrier Density', 'Location', 'B');
%}

%figure with currents on same plot
%PELTIER PLOT
%{
figure;
plot(xmid,q*J_E,'r-','LineWidth',2); hold on;
plot(xmid,q*J_n,'g-','LineWidth',3); hold on;
plot(xmid,q*(J_E + J_n),'k--','LineWidth',4);
plot(xmid,qJ_applied,'b--','LineWidth',2);
set(gca,'LineWidth',3); set(gca,'FontSize',20);
flux_range3 = max([abs(q*J_E) abs(q*J_n)]);
xlim([x(1) x(end)]);
ylim([-1.1*flux_range3 1.1*flux_range3]);
xlabel('Position (m)'); ylabel('Current Density (A m-2)');
title('Current Density By Source');
%}

%{
xlabelposE = 1e9*xmid(round(num_grid_points*0.6));
ylabelposE = q*1e-6*J_E(round(num_grid_points*0.6))*0.95;
text(xlabelposE, ylabelposE, '\uparrow J_E', 'VerticalAlignment', 'top', 'FontS
xlabelposn = 1e9*xmid(round(num_grid_points*0.08));
ylabelposn = q*1e-6*J_n(round(num_grid_points*0.08));
text(xlabelposn, ylabelposn, '\leftarrow J_{\nabla}', 'FontSize', 30, 'Color',

```

```

xlabelpostot = 1e9*xmid(round(num_grid_points*0.94));
ylabelpostot = q*1e-6*(J_E(round(num_grid_points*0.94))+J_n(round(num_grid_poin
text(xlabelpostot, ylabelpostot, 'J_{tot} \uparrow', 'VerticalAlignment', 'top'
%}

%legend('J_E', 'J_{\nabla n}', 'J_{\nabla T}', 'J_{total}', 'Location', 'B');

end

```

# Bibliography

- [1] A Majumdar. *Thermoelectricity in Semiconductor Nanostructures*. Science, Vol. 303, no. 5659, pp. 777-778. 6 Feb 2004.
- [2] K P Pipe. *Bipolar Thermoelectric Devices*. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Ph.D. Thesis. February 2004.
- [3] H B Callen. *Thermodynamics and an Introduction to Thermostatistics*. John Wiley and Sons, Inc., pp. 307-325. 1985.
- [4] G S Nolas, J Sharp, H J Goldsmid. *Thermo-electrics: Basic Principles and New Materials Developments*. Springer-Verlag. 2001.
- [5] G J Snyder and T S Ursell. *Thermoelectric Efficiency and Compatibility*. Physical Review Letters, Vol. 91, No. 14, pp. 148301-1-148301-4. 2 October 2003.
- [6] G J Snyder. *Application of the Compatibility Factor to the Design of Segmented and Cascaded Thermoelectric Generators*. Applied Physics Letters, Vol. 84, No. 13, pp. 2436-2438. 27 January 2004.
- [7] W Kim, et. al. *Thermal Conductivity Reduction and Thermoelectric Figure of Merit Increase by Embedding Nanoparticles in Crystalline Semiconductors*. Physical Review Letters, Vol. 96, No. 045901, pp. 1-4. 2 February 2006.
- [8] C Kittel and H Kroemer. *Thermal Physics, Second Edition*. W H Freeman and Company. 1980.
- [9] T E Humphrey and H Linke. *Reversible Thermoelectric Nanomaterials*. Physical Review Letters, Vol. 94, No. 096601, pp. 1-4. 9 March 2005.
- [10] D Vashaee and A Shakouri. *Improved Thermoelectric Power Factor in Metal-Based Superlattices*. Physical Review Letters, Vol. 92, No. 10, pp. 106103-1-106103-4. 11 March 2004.
- [11] D Vashaee and A Shakouri. *Electronic and Thermoelectric Transport in Semiconductor and Metallic Superlattices*. Journal of Applied Physics, Vol. 95, No. 3, pp. 1233-1245. 29 October 2003.

- [12] D Vashaee, Y Zhang, A Shakouri, G Zeng, and Y J Chiu. *Cross-Plane Seebeck Coefficient in Superlattice Structures in the Miniband Conduction Regime*. Physical Review B, Vol. 74, No. 195315, pp. 1-7. 9 November 2006.
- [13] W Nakwaski. *The Peltier Coefficient in Degenerate and Non-Degenerate Semiconductors*. Electron Technology, Vol. 14, Iss. 1, pp. 81-99. 1981.
- [14] G D Mahan. *Density Variations in Thermoelectrics*. Journal of Applied Physics, Vol. 87, No. 10, pp. 7326-7332. 9 February 2000.
- [15] L E Bell. *Cooling, Heating, Generating Power, and Recovering Waste Heat with Thermoelectric Systems*. Science, Vol. 321, pp. 1457-1461. 12 September 2008.
- [16] N S Hudak and G G Amatucci. *Small-Scale Energy Harvesting through Thermoelectric, Vibration, and Radiofrequency Power Conversion*. Journal of Applied Physics, Vol. 103, No. 101301, pp. 1-24. 20 May 2008.
- [17] D M Rowe, et. al. *CRC Handbook of Thermoelectrics*. CRC Press, LLC. 1995.
- [18] J Chen, Z Yan, G Lin, and B Andresen. *On the Curzon-Ahlborn Efficiency and its Connection with the Efficiencies of Real Heat Engines*. Energy Conversion and Management, Vol. 42, No. 2, p. 173. 9 August 2000.
- [19] H Yugami, H Sasa, and M Yamaguchi. *Thermophotovoltaic Systems for Civilian and Industrial Applications in Japan*. Semiconductor Science and Technology, Vol. 18, No. 5, p. 239. 4 April 2003.
- [20] P M Mayer. *High-Density Thermoelectric Power Generation and Nanoscale Thermal Metrology*. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, Ph.D. Thesis. April 2007.
- [21] J LaGrandeur, D Crane, and A Eder. *Vehicle Fuel Economy Improvement through Thermoelectric Waste Heat Recovery*. DEER Conference. Chicago, Illinois, USA. 25 August 2005.
- [22] U Ghoshal and R Schmidt. *Refridgeration Technologies for Sub-Ambient Temperature Operation of Computing Systems*. IEEE International Solid-State Circuits Conference 2000, 07803-5853-8/00. 2000.
- [23] M Lundstrom. *Fundamentals of Carrier Transport - Second Edition*. Cambridge University Press. 2000.
- [24] M G Kanatzidis, T Hogan, E Case, V Dravid, and H Schock. *Development and Application of High Efficiency Thermoelectric Materials for Power Generation*. 2008 ONR Direct Thermal-to-Electrical Energy Conversion Workshop. 9 December 2008.
- [25] J Cai and G D Mahan. *Effective Seebeck Coefficient for Semiconductors*. Physical Review B, Vol. 76, No. 075201, pp. 1-3. 2 August 2006.



- [26] S Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, pp. 37-43. 1995.
- [27] A Hochbaum, R Chen, et. al. *Enhanced Thermoelectric Performance of Rough Silicon Nanowires*. Nature, Vol. 451, pp. 163-167. 10 Jan 2008.
- [28] R Venkatasubramanian, E Siivola, T Colpitts, B O'Quinn. *Thin-film Thermoelectric Devices with High Room-Temperature Figures of Merit*. Nature, Vol. 413, pp. 597-602. 24 August 2001.
- [29] T C Harman, P J Taylor, M P Walsh, B E LaForge. *Quantum Dot Superlattice Thermoelectric Materials and Devices*. Science, Vol. 297, No. 5590, pp. 2229-2232. 27 September 2002.
- [30] I Avci, et. al. *Three-Dimensional TCAD Process and Device Simulations*. 16th Biennial University/Government/Industry Microelectronics Symposium, 2006.
- [31] M Lu, S Bowden, U Das, and R Birkmire. *a-Si/c-Si Heterojunction for Interdigitated Back Contact Solar Cell*. Proc. of 22nd European Photovoltaic Solar Energy Conf., 2007.
- [32] M Lu, S Bowden, and R Birkmire. *Two Dimensional Modeling of Interdigitated Back Contact Silicon Heterojunction Solar Cells*. International Conference on Numerical Simulation of Optoelectronic Devices, 2007.