High-density thermoelectric power generation and nanoscale thermal metrology

By

Peter M. Mayer

S.B. Electrical, Engineering Massachusetts Institute of Technology (1999)
S.B. Physics, Massachusetts Institute of Technology (2002)
M.Eng. Electrical Engineering, Massachusetts Institute of Technology (2002)

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

Doctor of Philosophy

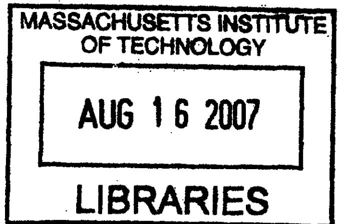
at the

MASSACHUSETTS INSTITUTE of TECHNOLOGY

APRIL 2007
[June 2007]

© 2007 Massachusetts Institute of Technology
All rights reserved

ARCHIVES



Author ______________________________
Department of Electrical Engineering and Computer Science
April 1, 2007

Certified by ______________________________
Rajeev J. Ram
Associate Professor of Electrical Engineering
Thesis Supervisor

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

High-density thermoelectric power generation and nanoscale thermal metrology

by

Peter M. Mayer

Submitted to the Department of Electrical Engineering and Computer Science
on April 1, 2007 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Electrical Engineering

ABSTRACT

Thermoelectric power generation has been around for over 50 years but has seen very little large scale implementation due to the inherently low efficiencies and powers available from known materials. Recent material advances appear to have improved the technology's prospects.

In this work we show that significantly increased generated power densities are possible even for established material technologies provided that parasitic losses are controlled and effective strategies are found for handling the large resulting heat fluxes. We optimize the performance of a thermoelectric generator in this regime, and discuss fundamental performance limits in this context. We present a design of a thermoelectric generator using conventional material and a microchannel heat sink that we predict can generate many times the power of a conventional thermoelectric, at a comparable efficiency. A high temperature vacuum test station is used to characterize the power generation, efficiency, and material properties of thermoelectric materials and generators. The results of a series of studies on various bulk and thin-film materials are presented, as well as packaged generator performance. The method of CCD thermoreflectance imaging is pursued in this thesis as a quantitative means for making non-contact temperature measurements on solid-state samples at the micro- and nano-scale. We develop and test a theory of the instrument and the measurement process to rigorously characterize the accuracy and precision of the resulting thermal images. We experimentally demonstrate sub-micron spatial resolution and sub-20 mK temperature resolution with this tool. High-resolution thermal images of thermoelectric elements, polysilicon-gate field effect transistors, and other integrated electronic devices are presented.

Thesis Supervisor: Rajeev J. Ram
Title: Associate Professor of Electrical Engineering

Acknowledgements

There are several people to whom I am greatly in debt for their help with the work presented in this thesis. Professor Rajeev Ram has been my advisor through both my MEng and PhD degrees here at MIT. His continual push for deeper understanding and his formidable command of a wide range of research areas have often moved me forward. His creativity, optimism, and integrity have at times sustained me. The further I came along in my PhD career, the more I understood how much I and other members of my group benefited from his support and leadership. He's had a profound influence on me in many ways which go beyond academic and technical matters.

My colleagues in my group, both past and present, have been the best part of my education and my time at MIT. Here I can mention only a few people, in no particular order. Farhan Rana was my first academic collaborator and a mentor when I joined the group. By example, he taught me what it was to fully analyze and understand a problem. Mathew Abraham was another early friend from the group. His rare combination of intense work ethic, kindness, and (very uncommon) sense is a source of inspiration for me. Harry Lee solves problems, just as a fire hydrant provides water. He helped me get started in the group and taught me much of what I know about measurement. Harry embodies what it is to be an engineer and to attack a problem without fear. Kevin Lee has taught me how to be fearless in other ways, including those typically addressed with life-insurance. Owing to his talent, his natural leadership ability, and his generous nature, he has taken on many thankless jobs that keep the lab running smoothly. Tauhid Zaman is now the senior grad student in the group. He has not let MIT claim his individuality, and he is an utterly unique coherent superposition of "gangsta" and "photonics geek." Xiaoyun Guo, our group's current post-doc, worked with me on the key problem of thermoelectric heat management. I have also benefited greatly from her expertise in preparing samples for thermoreflectance studies. Dietrich Lueerssen, formerly a post-doc at MIT and Mount Holyoke, has been a good friend and teacher. He brought CCD thermoreflectance imaging to the lab, wrote the software that enabled my thermoreflectance measurements, and allowed me to present in his stead at an important conference when I was starting up in the field. He pushed a patent through on which I was honored to be a co-author, and worked closely with me on my favorite publication from my PhD work. When he left to pursue other opportunities, he allowed me and others to reap the benefits of his hard work. Professor Janice Hudgings and Maryam Farzaneh of Mount Holyoke collaborated with me on the

thermoreflectance work. I enjoyed working with them both, and I look forward to seeing their future work in the field. Reja Amatya joined our group relatively recently and is continuing some of the thermoreflectance and thermal-themed work in the group. With enviable grace, she seems to have already mastered the things from my work that are applicable to her own. It's been fun to work with and hang out with her, but punishing to try and drink with her. Jason Orcutt, another relatively recent addition to the group, is a very talented engineer with refined tastes in coffee, beer, and burgers. Maybe other things too, but I wouldn't know. His warmth, wit and dry sense of humor helped sustain me during my thesis writing. Tom Liptay has been my closest friend for a very long time. As far as this thesis work goes, I've benefited from his incisive questions, his physical insight, and from his exacting (and evidently second nature) standard for total honesty in all things.

My work was financially supported by the Office of Naval Research and the National Science Foundation. I would like to thank Joshua Zide (UCSB), Gehong Zeng (UCSB), and Professor Ali Shakouri (UCSC) for the ErAs nanodot samples, Dr. Theodore Harman and Bob Reeder of Lincoln Labs for the n-QDSL samples, and Dr. Hylan Lyon, Jeff Sharp, and the team at Marlow Industries for providing the MAM material. Anita Villanueva was kind enough to provide the GaAs HEMT that appears in the future work section.

My wife Jen, my Mom and Dad, and my sisters and brothers have been my most important source of strength and wisdom. Jen's love is the source of much of the happiness in my life. To the extent that such things can be acknowledged in words, I do so here.

Table of Contents

Table of Contents 6
Table of Figures 8
Chapter 1: Introduction 16
1.1 Thermoelectric effects – physical intuition 16
1.2 Thermoelectric applications 20
1.3 Thesis Outline 30
Chapter 2: Theory and optimization of high-power density thermoelectric elements 32
2.1 Overview 32
2.2 Basic expressions for TEG 32
2.3 Optimum power or efficiency? 35
2.4 High Z thin-film TE materials – the heat sink limited regime[32] 39
2.5 Improved model for thin-film TE generator 40
2.6 Finite difference solver for Onsager relations 43
2.7 Optimum load in heat sink limited regime 45
2.8 Optimum length in heat sink limited regime 48
2.9 Optimum material 50
2.10 Conclusions 51
Chapter 3: High power density heat management 53
3.1 Overview 53
3.2 The importance of heat management 53
3.3 Heat spreading for thermoelectric power generation 56
3.4 Heat spreading in the transient regime 61
3.5 Experimental characterization of a heat sink 65
3.6 Microchannel heat sink – channel simulations 68
3.7 Modeling an integrated TE generator and microchannel cooler 76
3.8 Comparisons with the literature 93
3.9 Conclusions 101
Chapter 4: Thin-film TE performance characterization 103
4.1 Overview 103
4.2 Thermoelectric characterization techniques 103
4.3 High temperature power generation test benches 119

4.4 Bulk thermoelectric elements and generators....131
4.5 Thick film measurements....143
4.6 Thin film measurements - ErAs nanodot superlattice....151
4.7 Conclusions....161
Section 5: Nanoscale thermal imaging....162
5.1 Overview....162
5.2 Physics of thermoreflectance....162
5.3 Thermoreflectance imaging technique....168
5.4 Theory of sub-quantization thermoreflectance imaging....177
5.5 Experimental verification of sub-quantization imaging theory....184
5.6 Thermoreflectance calibration measurements....187
5.7 Polysilicon-gate field effect transistors....194
5.8 Cross-plane thermoreflectance imaging of Bi_2Te_3....202
5.9 Conclusions....210
Chapter 6: Conclusions and Future Work....212
Appendix 1: Justification of the standard linear treatment of the quantization process 227
Appendix 2: Thermoreflectance phase noise analysis....231
Appendix 3: Marlow MAM characterization data....233
Appendix 4: Thermal quadrupoles[55] used to estimate to meter bar errors....234
Appendix 5: Details of microchannel flow (calculation of Nu and f_D)....239
Appendix 6: MATLAB code....244
Chapter 2 code:....244
Finite difference solver for Onsager relations....244
Chapter 3 code:....256
Heat spreading and thermal quadrupole code....256
Microchannel channel simulations....260
TE generator + Microchannel system level simulations....264
Chapter 4 code:....277
Hi-Z module performance prediction....277
Chapter 5 code:....280
Function predicting errors of thermoreflectance measurements....280
Utility that is used to graphically select a region of a figure and return the correct $|\Delta R/R|$. 283
Function that corrects for the finite shutter time in the camera (~5% correction)....285
Procedure used for modeling of the dynamic temperature distribution in MAM material... 286
1-D finite difference time domain Crank-Nicolson solver....289
1-D finite difference time domain Crank-Nicolson solver with arbitrary time and space step-sizes....293
Bibliography....299

Table of Figures

Figure 1. A schematic depiction of the Seebeck effect in a box of gas. 17

Figure 2. A schematic depiction of the Peltier effect in an n-type semiconductor between two metal contacts. 19

Figure 3. A power generating TE couple, consisting of an n-type element wired electrically in series and thermally in parallel with a p-type element. 20

Figure 4. RTG used in Cassini space probe (p532, CRC Handbook of Thermoelectrics). 21

Figure 5. Close-up of GPHS unicouple (from http://www.its.caltech.edu/~jsnyder/thermoelectrics/segunicouple_page.htm). 22

Figure 6. "Thermic" thermoelectric wristwatch by Seiko (from Seiko website). 23

Figure 7. Thermoelectric elements in the Thermic wristwatch (from Seiko website). 23

Figure 8. ZT vs. mean operating temperature for several n-type (left) and p-type (right) bulk thermoelectric materials. (Images from Jeff Snyder, http://www.its.caltech.edu/~jsnyder/thermoelectrics/) 24

Figure 9. Thermoelectric generator and heat exchanger envisioned for the BMW 5-series (from http://www.greencarcongress.com/2006/09/bmw_introduces_.html). 25

Figure 10. Recent advances in thin-film thermoelectrics. (from Majumdar, 2004) 27

Figure 11. SiGe superlattice thermoelectric material. (from Fan et. al., 2001) 28

Figure 12. Infrared image of a computer chip fault during real conditions testing using the FLIR Systems ThermaCAM (http://www.flirthermography.com/success/ir_image/1030/industry_id/1000/). 29

Figure 13. Schematic diagram of a thermoelectric generator. 33

Figure 14. Power vs. efficiency curves for gas turbine engine with varying internal frictional losses, demonstrating general features of heat engines with internal irreversibilities. 36

Figure 15. Normalized power vs. efficiency curves for TE elements with various figures of merit. The rightmost curve allows Z to approach infinity, corresponding to a perfect material with no internal losses. The curves are generated by varying the external load resistance. Note that for typical Z's, the optimum efficiency and optimum power points are close to one another. 37

Figure 16. Additional parasitics due to packaging that can be easily included in the model using an effective contact resistance and an effective heat sink parameter. 42

Figure 17. Finite difference algorithm used to compute 1-D transport including arbitrary grading and temperature dependence of material parameters. 44

Figure 18. Optimal impedance matching when heat sink limited. 48

Figure 19. Power density versus leg length for different heat sinks and improvements in selected material parameters. 49

Figure 20. Effects of material improvements on generator performance. The first figure shows the change in power vs. efficiency loops due to increasing in Z by a factor of two for using each of the three material parameters. The second figure extends the analysis to larger changes in Z. 50

Figure 21. HZ-2 thermoelectric module from Hi-Z Technology, Inc. 54

Figure 22. Geometry and parameters used for heat spreading calculations. A concentric cylindrical geometry is assumed to allow for a closed form solution in terms of orthogonal functions. 58

Figure 23. Finite element solution for the temperature distribution at the top of a heat spreader using FEMLAB. The color scale is proportional to temperature above room temperature. The isotherms are nearly circular, even though the heat source is square-shaped. 59

Figure 24. The effect of heat spreading on the temperature difference across a thin-film element and on the generated power. 60

Figure 25. The equivalent thermal impedance network used to model the heat spreader. A thermal resistance modeling the heat sink on the bottom of the heat spreader has also been included. 63

Figure 26. A comparison of various analytic models for the cylindrical thermal spreading resistance. The dotted 'steady-state' line models the steady state thermal resistance model of Lee, Song, Van Au, and Moran discussed in the previous section. The 'half-space' line is the solution for the dynamic spreading resistance into an infinite half-space. The 'quadrupole' line is the solution of the dynamic quadrupole model presented in this section. 64

Figure 27. Basic geometry for a measure bar used to measure the heat transfer coefficient into a heat sink. 66

Figure 28. Experimental determination of the heat transfer coefficient for a chilled water cold plate. Raw thermocouple data is shown at left, and the inferred heat transfer coefficient is shown at right. 67

Figure 29. Schematic diagram of microchannel cooler geometry (from Phillips, 1990). 69

Figure 30. Domain of microchannel simulation. The parameters a and b are the channel wall lengths. Fluid flow is in the x-direction. 73

Figure 31. Schematic diagram of TE generator with an idealized heat sink showing important dimensions. 78

Figure 32. Generated power density and efficiency for benchmark TE thin-film generator, as a function of area coverage. 80

Figure 33. Layer temperatures for a TE generator (benchmark heat sink) as the coverage area of the TE elements on the spreader is varied 81

Figure 34. Generated power density and efficiency as the length of the thermoelectric element is varied. 82

Figure 35. Schematic picture of TE generator integrated with microchannel cooler, showing important dimensions. 82

Figure 36. Side view of microchannel cooler showing one of the repeating blocks of the generator (one TE element with its corresponding heat sink area). Also shown are the temperatures used to define the thermal resistances for the heat transport model. 84

Figure 37. Efficiency (solid line), generated power (dash-dot line), and net generated power density (dotted line) for varying mean water flow velocities. Net generated power is defined as generated power minus the power needed to pump the water through the channel.................. 87

Figure 38. Nusselt number and Darcy friction factor as a function of Reynolds number. Note the effects of turbulence, modeled here at Re>2300. 89

Figure 39. Efficiency (solid line), generated power density (dash-dot line), and net generated power density (dotted line) for various microchannel aspect ratios. Flow velocity is held constant 89

Figure 40. The effects of varying aspect ratio on the channel heat sink parameter h and on the effective heat sink parameter heff, which includes the area enhancement from the fins. 90

Figure 41. Performance of the microchannel cooled thermoelectric system as a function of the amount of heat spreading used. The optimum net power density is at an area ratio of unity, meaning that heat spreading is not helpful here. 91

Figure 42. For a shorter (5 micron thick rather than 20 micron thick) thermoelectric element, heat spreading may make sense. 91

Figure 43. The generated power density and efficiency of the thermoelectric microchannel generator as the length of the thermoelectric elements are varied. The efficiency becomes negative at long element lengths due to the fixed power lost pumping the microchannels. 92

Figure 44. Power density vs. efficiency, with the variation of TE element length used as a parameter. The dotted curve is the TE generator with the ideal passive heat sink, and the solid curve is the microchannel actively cooled generator. 93

Figure 45. Comparison between the measurements of Tuckerman, Pease (1981) and the microchannel model developed for this thesis. The maximum thermal resistance between the surface of the microchannel heat sink and the inlet water is plotted as a function of channel water flow rate. 95

Figure 46. Breakdown of the thermal resistances of our theoretical calculation of the Tuckerman and Pease thermal resistance. 96

Figure 47. The percent change in the thermophysical properties of water from their room temperature values as temperature is varied. 97

Figure 48. Comparison of microchannel simulation code developed for this thesis and Figure 4 (b) from Qu and Mudawar (2002). The symbols correspond to Qu and Mudawar's measurements and also to their theoretical predictions of the pressure drop across the microchannels for two different input heat powers (100 W/cm^2 and 200 W/cm^2). The solid line depicts the prediction of out theory for 100 W/cm^2 of input power. 99

Figure 49. A comparison of the microchannel simulation code developed for this thesis to the results of Qu, Mudawar (2002). The plot depicts the total temperature change in the water flowing through the microchannels betwen the inlet (T_{in}) and outlet (T_{out}). The symbols (measurements) and solid lines (theory) are from Qu and Mudawar, and correspond to two different levels of input heat flux (100 W/cm^2 and 200 W/cm^2). The dotted line is our theoretical prediction corresponding to an input heat flux of 100 W/cm^2. 100

Figure 50. A comparison between the microchannel heat sink theory developed for this thesis and the results of Qu and Mudawar (2002). The symbols depict the experimental temperature measurements (T_{th}) of Qu and Mudawar from four thermocouple probes positioned in the substrate 2.462 mm above the top of the microchannels relative to the inlet fluid temperature T_{in}). The measurement was for an input heat flux of 100 W/cm^2. Location 1 corresponds to a point 5 mm from the inlet port, and locations 2, 3, and 4 are spaced 11.588 mm from one another along the channel. The solid line shows our theoretical prediction for the average temperature difference between the top of the microchannel and the inlet fluid temperature. 101

Figure 51. The thermoelectric response of a 10 micron thin film on a 90 micron substrate during a simulated transient Harman measurement. The figure on the right is a zooms in on the region of time immediately after the current pulse is turned off. 107

Figure 52. Important parasitic resistances for conventional transient Harman measurement and thin-film transient Harman measurement. 108

Figure 53. A plot of the two contact thermal resistances present during a thin-film transient Harman measurement, as well as the thermal resistance of the TE element. 109

Figure 54. Thermal conductivity test-fixture (Tzeng, 2000) 111

Figure 55. Circuit used to perform 3ω measurement. To measure $V_{3\omega}$, the voltage across the device (DUT) is subtracted from a linear reference voltage to increase the avalable dynamic range of the lock-in. 114

Figure 56. Example of 3ω measurement, performed on a platinum resistive heater suspended on a silicon membrane (see inset). The frequency dependence of the temperature resembles a single-pole transfer function with a pole corresponding to the thermal capacitance of the silicon membrane. 114

Figure 57. 2-D time domain finite difference simulation of a Van der Pauw in-plane electrical conductivity measurement on a suspended thin-film. The probe current can significantly perturb the sample temperature, adding a spurious thermal voltage to the measured Ohmic voltage drop. 115

Figure 58. Sample geometry for in-plane Seebeck coefficient determination. 117

Figure 59. A schematic of a cross-plane Seebeck coefficient measurement (Zeng et al., 2005) 118

Figure 60. Experimental setup for measuring generated power density. 120

Figure 61. Active load used for the measurement of the device current and voltage. In preliminary measurements, an initial resistance calibration was used to dispense with the precision current sensing resistor. 123

Figure 62. I-V curves of active load, taken with semiconductor parameter analyzer. A parasitic series lead resistance of around 20 mOhms is also included in the circuit. 123

Figure 63. Milli-Ohm active load with linearizing resistors. 124

Figure 64. A test of the linearized active load for mOhm load synthesis. For a fixed gate voltage, the current through the active load was varied over several decades with no significant change in load resistance observed. 125

Figure 65. The high-temperature vacuum thermoelectric test bench (a.k.a. Shrek). 126

Figure 66. Schematic diagram of the thermal measurements inside of the high-temperature vacuum test bench. 127

Figure 67. Example of the raw data from a measurement on a thermoelectric generator using the high-temperature vacuum system. 129

Figure 68. Comparison of the open circuit voltages for two samples measured using the first generation (rapid power measurement) and second generation (vacuum power and efficiency measurement) setups. 130

Figure 69. A comparison of different interface thermal resistances used to contact thermoelectric elements. 130

Figure 70. Top photo is a traditional Bridgeman-grown thermoelectric element. Lower photo is the Marlow micro-alloy material (MAM). Photos taken from Marlow website (http://www.marlow.com/AboutMarlow/advanced_materials.htm). 132

Figure 71. Experimental optimization of the load resistance. 134

Figure 72. Load resistance vs. length for Bi_2Te_3 elements 134

Figure 73. Generated power and open circuit voltage for Bi_2Te_3 element. 135

Figure 74. Open circuit voltage vs. length for Bi_2Te_3 elements. 136

Figure 75. Generated power density vs. length for Bi_2Te_3 elements. 138

Figure 76. Thermal conductivity and Seebeck coefficient data from the Marlow MAM calibration sample. The circles are Marlow's measurements, and the x's are the measurements made using the high-temperature vacuum setup built for this thesis. 139

Figure 77. The top chart compares the measured element material parameters to book material values. The bottom chart compares the predicted performance of the generator based on the measured material against the stated performance of the generator from the manufacturer's data sheet. 142

Figure 78. The geometry of gold-QDSL couple used by Harman et al. 147

Figure 79. A > 1 day-long measurment of the Seebeck coefficient and the thermal conductivity of the Lincoln Labs G-207 (n-QDSL) sample. This long measurement was done in air to prolong the life of the device. 151

Figure 80. Cross-sectional transmission electron micrograph of ErAs nanodot material, confirming the randomly spaced nanometer-sized ErAs island formation. 153

Figure 81. Load and generated power vs. active load control voltage. Inset shows schematic of test setup. 155

Figure 82. Generated power densities for ErAs superlattice material on differing substrate thicknesses, as a function of applied temperature difference. Inset shows open circuit voltages over same temperature range. 156

Figure 83. Measured and expected open circuit voltages for different substrate lengths. The expected contribution from the superlattice and the substrate are shown. 157

Figure 84. Slope of open circuit voltages with respect to the hot side temperature for ErAs superlattice material and for a representative Bi_2Te_3 sample. 158

Figure 85. UCSB 400 element thermoelectric generator (photo courtesy of Gehong Zeng).... 160

Figure 86. Thermoreflectance spectrum of silicon. (Miyazaki, 1995) 164

Figure 87. Theoretical calculations of the changes in the index of refraction expected in InP for three doping levels relative to an undoped sample. Three contributing factors (free carrier, band filling, and bandgap shrinkage) cause the changes. (Bennett et al.,1990) 167

Figure 88. Timing diagram for a measurement of Ohmic (I^2R) heating using the four-bucket technique. The frequency of the heating occurs at 2x the frequency of the incident current, and the camera is triggered at 4x the frequency of the temperature oscillation. The four resulting images are accumulated into bins labeled here as I_1, I_2, I_3, and I_4 for post-processing. 169

Figure 89. The basic experimental setup used for the thermoreflectance studies of this work. 170

Figure 90. Measurement below the quantization threshold using oversampling and averaging. With no noise (top), the dynamic range of the measurement is limited by the quantization step size, regardless of any averaging. With sufficient noise (bottom), the dynamic range can be improved through averaging, allowing a signal between the quantization levels to be measured. 175

Figure 91. Results using CCD thermoreflectance imaging from Grauby et al. (1999, 2001).... 176

Figure 92. A thermoreflectance phase image illustrating partical phase convergence. The solid area depicts a polysilicon resistive snake in a Si substrate. The speckled signal elsewhere is a result of random phase due to insufficient signal power........ 179

Figure 93. A thermoreflectance measurement for varying numbers of iterations. The dashed line is the ideal thermoreflectance whose measurement is desired. The solid line describes the expected value (first moment) of the thermoreflectance. The dotted lines describe the +/- 1 sigma error bars of the measurement........ 183

Figure 94. Example of plots of ε_{offset} and ε_{error} as a function of iterations and CCD counts. The plots here allow the error to be predicted given a value α=0.207 and β=7.650 for the fundamental noise in the camera, and a thermoreflectance signal of magnitude $\Delta R/R=10^{-4}$... 184

Figure 95. Experimental confirmation of sub-quantization level imaging. Thermal images (top inset) of a boron diffused resistor in Si were used to find the temperature profile (lower inset) across the resistor. The average temperature inside the resistor obeyed the expected linear relationship with electrical power. The temperature corresponding to the quantization limit was 2.93 K; all data points are below that........ 185

Figure 96. Example of plots of ε_{offset} and ε_{error} as a function of iterations and CCD counts. The plots here allow the error to be predicted given a value α=0.207 and β=7.650 for the fundamental noise in the camera, and a thermoreflectance signal of magnitude $\Delta R/R=10^{-4}$. The measured signal and measured +/- 1 sigma error bars are shown in gray. A black solid line depicts the theoretical signal, and the black dotted lines depict the theoretical +/- 1 sigma error limits........ 186

Figure 97. Thermoreflectance calibration procedure used in this work. A thermoreflectance measurement on a uniform area of the sample is performed, while on an adjacent area, a microthermocouple monitors the surface temperature. A Peltier stage is driven with a sinusoidal current to modulate the sample temperature. 187

Figure 98. Example of the simultaneous acquisition of three thermoreflectance coefficients (calibrations) using one thermoreflectance image and a microthermocouple (not shown). The top photo shows a test structure on a Silicon wafer and the three areas used. The bottom picture is the phase image of the thermoreflectance measurement. 189

Figure 99. The left-most panel shows a comparison of temperature measured in diffused resistor using a thermorelflectance measurement and using the 3ω technique. The mean thermoreflectance temperature in the resistor is marked with the dotted line and the shaded gray region is the 1-sigma confidence interval for the 3ω temperature. The right panel shows the measured temperature vs bias for the 3ω measurement........ 191

Figure 100. An example of an imperfect A-to-D quantizer. The presence of unequal bin widths simulates the quantizer's nonlinearity error. In this case, a differential nonlinearity (DNL) of 0.35 quantization steps has been assumed. The top inset shows a plot of the DNL error over 256 bins of the A-to-D converter, and the bottom shows a similar plot of the INL error........ 193

Figure 101. The image at left depicts the experimental setup for the FET thermal measurements. The image at right is a photomicrograph of the wafer with the FETs and other test structures. 194

Figure 102. The left-most image is a photo-micrograph showing the p-channel FET to be studies using thermoreflectance, with the the region of interest marked. The I-V characteristics of the FET are shown in the plots on the right. The different curves are for varying gate voltage..... 195

Figure 103. Calibrated thermal images of the FET in the saturation and linear bias regimes. The plots at right show the vertically averaged temperature profiles in the gate (drain is to the left of the gate in the images and profile) 196

Figure 104. Thermal image of a 5 micron gate-length FET. 198

Figure 105. Thermal image of a point defect in the gate of a FET. The defect was not evident in the photomicrograph (shown at left) but was readily apparent in the thermal image and the electrical response of the device. 199

Figure 106. Thermal image of the dynamic heating in a FET gate. The R-C losses in the gate as it is rapdly charged and discharged are responsible for the heating. 200

Figure 107. Schematic diagram of the thermoreflectance imaging apparatus used for measurement of thermoelectric elements (left) and calibration of the material thermoreflectance coefficient (right). 203

Figure 108. Phase image from Peltier measurement of Bi_2Te_3 sample. The speckle in the phase image is not visible in the photomicrograph of the sample, and the size of the speckle matches the expected polycrystalline grain size (~10-20 μm). 206

Figure 109. Cross-plane thermoreflectance images of Bi_2Te_3. The left most images depict the Peltier measurement (solid) and steady-state theory (dotted), and the right depicts the Ohmic. These measurements were made using thin copper thermal and electrical contacts and a 2A sinusoidal current at 0.5 Hz. 207

Figure 110. Peltier (left) and Ohmic (right) images demonstrating the importance of the frequency of excitation. These measurements were made using thin copper thermal and electrical contacts and a 2A sinusoidal current at 7 Hz. 208

Figure 111. The left plot plots the efficiency optimized for power (solid) and the efficiency optimized for efficiency (dashed) as a function of ZT. The right plot does the same for generated power density (optimized for power = solid, optimized for efficiency = dashed) for three element lengths (2mm, 400 μm, 80 μm) 213

Figure 112. The Peltier heating/cooling and the Joule heating powers across a 5 nm tunneling barrier (Lake et al. 1992). Some of the transport energy of the hot electrons goes into the (top image, antisymmetric, reversible) Peltier effect, and some goes into Joule heating (bottom image, symmetric, irreversible). Two biases ($\Delta\mu$=9 mV, 18 mV) are checked. 218

Figure 113. Image of contact heating in the drain of a pentacene channel organic FET device. (device courtesy of John Kymissis, Bulovic group, MIT) 221

Figure 114. Heating in a GaAs HEMT. (device courtesy of Anita Villanueva, del Alamo group, MIT) 222

Figure 115. Photomicrograph of the cross-plane view of an ErAs nanodot sample. The 4.9 μm superlattice is grown atop a 175 μm InP substrate. 223

Figure 116. Cross-plane thermoreflectance images of ErAs nanodot thermoelectric material on InP substrate. The current excitation was a 5 A amplitude, 5.5 Hz sine wave. 224

Figure 117. Cross-plane thermoreflectance on a 4.9 micron ErAs nanodot thin-film on a 180 micron InP substrate. 8 A pk-pk current producing a heating frequency of 3 Hz was was used to generate the temperature profiles. 225

Figure 118. Quantization process. The quantizer effectively rounds off the input signal w to the nearest quantization step. 227

Figure 119. The worst case dependence of the correlation between the quantization error and the signal on the input noise level. The inset shows the correlation as a function of the input noise level and the position of the signal in the quantization bin (x). 229

Figure 120. The second moment of the quantization error as a function of the input noise level (solid line). Also shown is the moment computed using the standard approximation σ_d^2 +1/12 (dotted line). 232

Figure 121. Cylindrical geometry considered for time dependent one dimensional heat transfer. 234

Figure 122. Impedance "T" description of passive thermal two-port. 236

Figure 123. Impedance model of copper rod terminated by a heat sink. 237

Figure 124. Thermal impedance of two heat-sunk copper bars of differing cross-sectional area, plotted as a function of frequency. The percent difference between each thermal impedance and its steady-state value is also plotted. The analysis is useful for determining the amount of time needed to perform meter bar heat flux measurements. 238

Chapter 1: Introduction

1.1 Thermoelectric effects – physical intuition

Thermoelectric (TE) generators rely on the joint Seebeck, Peltier, and Thomson effects to convert heat into electrical power. The Seebeck effect, first observed in 1821 in metals, refers to the voltage that develops when a thermal gradient is applied to a solid-state sample. The temperature dependent Seebeck coefficient $\alpha(T)$ with units of [Volts/Kelvin] is the differential voltage that is developed across a sample for a small applied temperature gradient with no current flowing (open circuit conditions). More fundamentally, it is the rate of change of the Fermi level of a solid with temperature divided by the charge of an electron $\left(\frac{1}{q}\frac{\partial f(T)}{\partial T}\right)$. It can also be regarded as the entropy transported per unit charge in a material[1]. The Peltier effect was discovered independently thirteen years later. Its effect is observable as local heating or cooling at a junction between two dissimilar materials when current is passed through the junction. An important characteristic of the effect is that it is reversible, so that when the direction of the current through the junction is switched, so is the sign of the heating/cooling. The difference in temperature dependent Peltier coefficients between two materials $\Pi_1(T)$-$\Pi_2(T)$, with units of [Volts=Watts/Amp], describes the amount of cooling or heating at the junction per unit current when both sides of the junction are at the same temperature. More fundamentally, the Peltier coefficient can be understood as the energy transported per unit charge measured relative to the Fermi level in the material. In 1858, by assuming that the Seebeck and the Peltier effects are reversible thermodynamic processes, Lord Kelvin (William Thomson) established that there must be a simple thermodynamic relationship between the two parameters such that $\Pi=\alpha T$. He also demonstrated the existence of a third parameter (the Thomson coefficient τ) which was needed to explain the reversible heat generated (or absorbed) within a homogenous material with an applied temperature gradient through which a current was passed. The Thomson coefficient can be more simply understood as the specific heat per unit charge in a material, or as the derivative of the material Seebeck coefficient with respect to absolute temperature. All three coefficients are useful in understanding the operation of practical TE devices. While Kelvin related the three coefficients on basic thermodynamic grounds, within the framework of modern statistical mechanics and solid state physics the

coefficients are probably best understood from the Boltzmann transport model for linearized electron transport[2].

A qualitative way to visualize the microscopic physics responsible for the Seebeck coefficient is to imagine that the free carriers in the solid-state TE material behave as a gas. If a temperature gradient is applied to a box of gas at constant pressure, the density of the gas throughout the box will change accordingly. On the cold side of the box the density of the gas will be higher, and on the hot side the density will be lower (Figure 1).

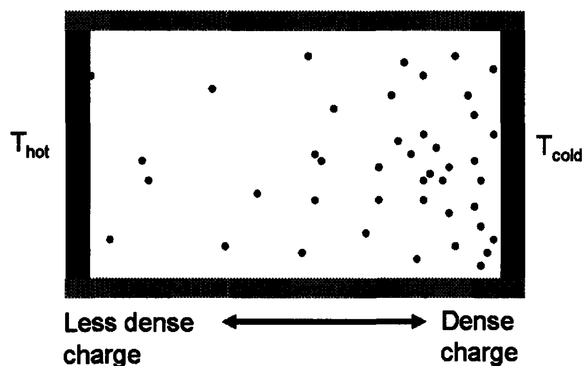


Figure 1. A schematic depiction of the Seebeck effect in a box of gas.

Since the free carriers in TE material are charged, the density variation gives rise to a voltage difference between the hot and cold sides. One can say the same thing for a realistic TE material using the language of solid-state physics and statistical mechanics: the temperature gradient in the solid (established mainly by the lattice and electrical thermal conductivities and the electron-phonon interactions which thermally couple the electron and phonon populations) gives rise to a spatial variation of the electronic carrier population, which can be described by a gradient in the carrier Fermi level, resulting in a voltage difference.

In semiconductors, both electrons and holes can contribute to the Seebeck coefficient, but it should be clear from the simple picture in Figure 1 that they typically do so with opposite signs. In a doped (extrinsic) semiconductor or semimetal, the majority carriers can be made to dominate the thermoelectric response of the material. The resulting Seebeck coefficient is therefore dependent on temperature and doping. For example, at high enough temperatures the Seebeck coefficient of semiconductors and semimetals will typically decrease as thermally

excited minority carriers begin to cancel out the Seebeck effect of the majority carriers. This is one reason for the prevalence of smaller band-gap thermoelectric materials at lower operation temperatures (e.g. Bi_2Te_3) and larger bandgap materials at higher operation temperatures (e.g. SiGe).

At room temperature, semiconductors and semimetals doped with the carrier concentrations commonly found in useful TE devices typically have Seebeck voltages in the neighborhood of 200 μV/K. The magnitude of the Seebeck voltage in metals is typically considerably smaller (~ 5 μV/K). This can be most directly seen from the fact that the electrons in a metal (or heavily doped semiconductor) are a degenerate Fermi gas, so that the available states for electron transport are tightly concentrated and nearly symmetrically distributed around the Fermi level, so that an applied temperature gradient has little effect on the electron energy distribution. Equivalently, the entropy change per added electron in a degenerate Fermi gas is much smaller than a classical gas. (This is actually a familiar fact that can be related to the lower molar specific heat in a degenerate Fermi gas relative to its classical counterpart.) With this picture we can understand why there is an optimal doping for a practical thermoelectric semiconductor, and why there is some tradeoff between electrical conductivity and Seebeck coefficient: some doping is necessary to overcome the intrinsic regime, and improves both the Seebeck effect and the electrical conductivity, but too much doping pushes the semiconductor into more metallic behavior, with still higher electrical conductivity but very low Seebeck coefficient.

It is worth discussing the Peltier effect in more detail. Consider a uniform n-type semiconductor with two metal contacts (Figure 2). For simplicity, the n-type semiconductor is assumed to be heavily doped so that the effects of conduction band bending on transport near the interface can be ignored. The materials may be assumed to be near thermodynamic equilibrium so that a constant Fermi level can be drawn. In spite of this, there is nothing requiring that the average energy associated with a transport electron in each material be identical. Both the potential energy of the electrons (determined by the band structure) and the kinetic energy of the electrons (their distribution within the allowed energy levels for transport) can be quite different in the different materials[3]. In the example of Figure 2, the carriers responsible for transport in the metal are closely packed around the Fermi level, while those in the n-type semiconductor are located in the conduction band. The band offset between the metal and in the semiconductor requires that electrons entering the n-type semiconductor from the metal obtain some energy to surmount the barrier ΔE_{Pelt}. This is provided by the local thermal energy of the

lattice, resulting in cooling at that junction. When the electrons exit the semiconductor, within a short distance of the junction (the energy relaxation length) they give up their extra energy to the lattice, resulting in heating at the other junction. The difference in mean transport energy per electron is exactly the difference between Peltier coefficients in the materials, in this case given by $\Pi \cong \Delta E_{Pelt}/q$ where q is the charge of an electron.

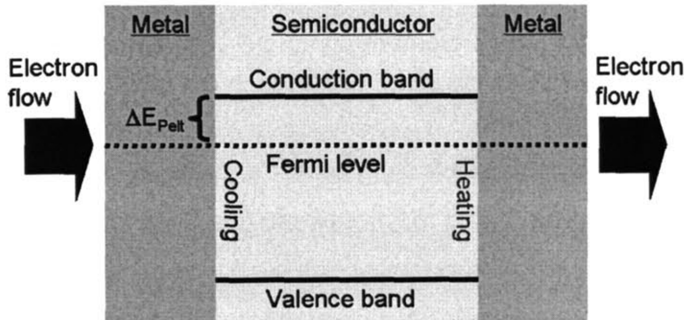


Figure 2. A schematic depiction of the Peltier effect in an n-type semiconductor between two metal contacts.

Semiconductor heterojunctions and homojunctions also exhibit Peltier effects, based again on the difference in mean transport energies in the materials [4]. Metal-to-metal junctions also exhibit a Peltier effect, although it is typically much smaller, since the electrons active in transport in most metals have very similar distributions in energy.

If a temperature gradient is applied to a thermoelectric material, the Seebeck effect can create a voltage across (and current through) some external load, generating useful electrical power. By connecting an n-type and a p-type thermoelectric element electrically in series but thermally in parallel (Figure 3), the resulting Seebeck voltages add with the same sign, and if several such couples are cascaded, large voltages can be generated from modest temperature differences. The thermoelectric generator can also be "run in reverse," that is, an applied current can be used to pump heat across the elements using the Peltier effect. To increase the amount of heat carried across the device per electron pushed through the circuit, the same trick of cascading TE couples is used. Commercial Peltier coolers based on this idea are used to thermally stabilize temperature sensitive electronic and optoelectronic components, and for small scale refrigeration (and heating) more generally.

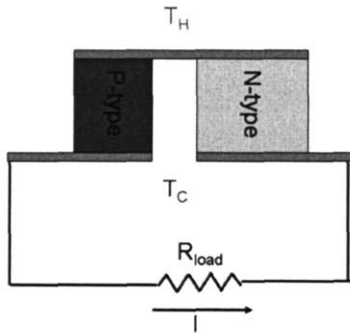


Figure 3. A power generating TE couple, consisting of an n-type element wired electrically in series and thermally in parallel with a p-type element.

It is interesting that in discussing the operation of the thermoelectric generator and the thermoelectric (Peltier) cooler it is convenient to switch back and forth between the Seebeck and Peltier view of things even though the microscopic phenomena responsible for each are intimately connected through Kelvin's thermodynamically derived (but somewhat physically opaque) relationships. After Kelvin's work Onsager subsequently showed that the simple relationship between the Peltier and Seebeck coefficients was actually a specific instance of a deep general relationship derived from time reversal symmetry applied to linearized transport equations. In a typical modern formulation of the theory one uses the Boltzmann transport formalism under the relaxation time approximation to write equations for the heat and charge flow in the material due to applied thermal and electrical fields. In this picture, the Peltier and Seebeck coefficients are derived transport coefficients computed with knowledge of the material-specific density of states and scattering mechanisms. When the problem is formulated in this way, the relationship between the Peltier and Seebeck coefficients results from the fact that the same electron distribution function f(k) is used to obtain the transport coefficients for both charge and heat.

1.2 Thermoelectric applications

To date, TE generators have been primarily used in space power systems and for off-grid terrestrial power in remote locations. An example of a radioisotope thermoelectric generator (RTG) that was used aboard NASA's Cassini space probe[5] (launched on 10/15/97) is shown in Figure 4 to illustrate some of the basic components of a TE generation system.

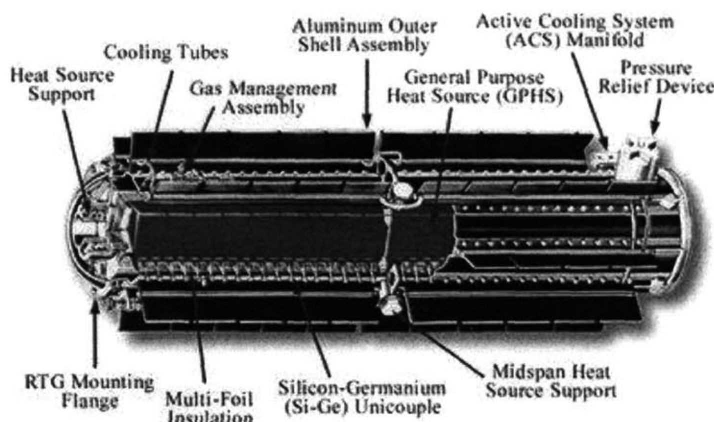


Figure 4. RTG used in Cassini space probe (p532, CRC Handbook of Thermoelectrics).

Like all thermoelectric generation systems, the Cassini RTG consists of a source of heat, the TE couples themselves, and a heat sink rejecting waste heat from the cold side of the generator. The unit is cylindrical, with a diameter of 42.2 cm, a length of 114 cm, and a weight of 55.9 kg. The hot side of the generator is powered from the 10.9 kg of radioactive ^{238}Pu in the core of the RTG, emitting approximately 4400 W of thermal power radially through the TE couples. The 572 TE couples are connected electrically in series and thermally in parallel. This is a common configuration in TE generators since the individual couples generate small voltages, in this case on the order of tens of mV per couple. The material used for the TE couples (shown in Figure 5) is a carefully chosen SiGe alloy capable of functioning at a hot side temperature of 1000 C with reasonable efficiency. The heat-flux through the 3 cm long couples generates about 300 W of electrical power at 30 V with 7% heat-to-electrical power conversion efficiency. An active pumped gas heat exchanger keeps the cold side of the generator at around 300 C and moves the waste heat to aluminum radiator fins. The heated gas is also used (cogeneratively) to control the temperature of other onboard systems. Similar RTGs were also used on the Voyager, Galileo, and Ulysses flights as a source of power for missions to the outer solar system where solar power was insufficient for the probes' electrical power requirements. Because the couples in TE generators have no moving parts they have a well-deserved reputation for reliability. In the case of the NASA's RTGs the lifespan so far has been limited not by mechanical or chemical failure of the elements but rather by the half-life (87.7 years) of the chosen radioisotope heat source.

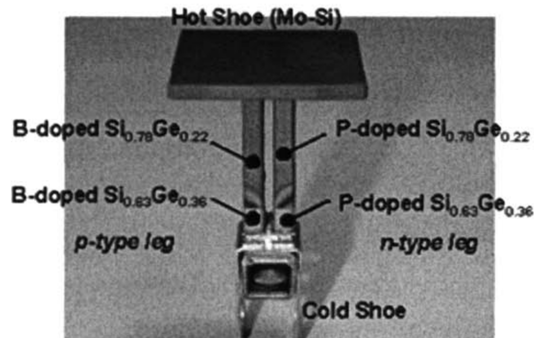


Figure 5. Close-up of GPHS unicouple (from http://www.its.caltech.edu/~jsnyder/thermoelectrics/segunicouple_page.htm).

While thermoelectric generation has been used for over 20 space missions since the 1960s, widespread commercial development of the technology has been hampered by the small efficiencies (~5%) attainable using typical bulk thermoelectric materials. Thermal efficiencies for modern large-scale power production facilities world-wide range between 30%-50% depending on the technology employed. The most likely commercial uses of thermoelectric generation in the near term are therefore not for large on-grid power generation, but rather for small off-grid sites and portable power applications.

An ambitious early commercial applications of microscale TE generators for a portable power application was Seiko's Thermic watch, introduced in December 1998[6]. The watch's basic design is shown in Figure 6, and an electron microscope image of the thermoelectric generator elements is shown in Figure 7. While the watch evidently functioned well, the costs involved in its production priced the watch out of the range of the target consumer market. After a limited production run the project was abandoned. However, several more recent research programs have been directed towards functional micro-thermoelectric generators and micro-thermoelectric fabrication techniques[7].

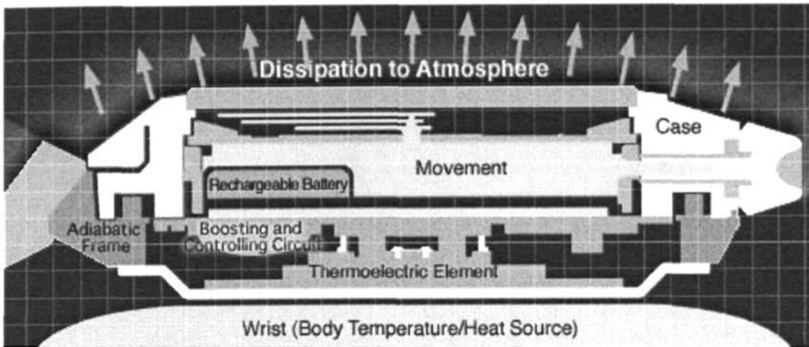


Figure 6. "Thermic" thermoelectric wristwatch by Seiko (from Seiko website).

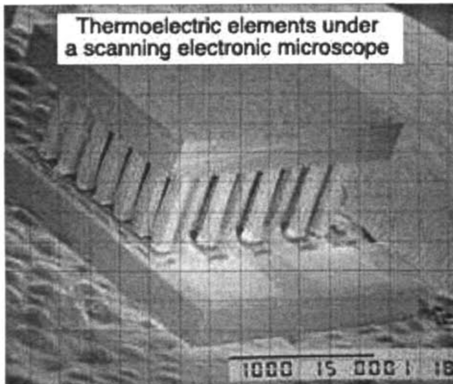


Figure 7. Thermoelectric elements in the Thermic wristwatch (from Seiko website).

Generator performance is described with two metrics, the thermodynamic efficiency (the ratio of power generated to power drawn from the heat source) and the power density (the power generated per unit area of the device). These metrics depend on the materials used for the TE couple, the temperatures of the hot and cold sides of the generator, and the load driven by the generator. Like all heat engines thermoelectric power generation is ultimately limited by the Carnot efficiency $\eta_C = \frac{T_h - T_c}{T_h}$. Practically speaking the maximum efficiencies for TE generation fall well below the Carnot limit and will be shown in Chapter 2.2 to be

$$\eta_\eta = \eta_C \frac{\sqrt{1+ZT}-1}{\sqrt{1+ZT}+T_c/T_h}, \qquad (1.2.1)$$

where $T = \frac{T_h + T_c}{2}$. Here ZT is the so-called thermoelectric figure of merit, a dimensionless parameter commonly used to evaluate a material's suitability for thermoelectric applications. Recent advances in TE generation have primarily focused on increasing the material ZT

because the maximum efficiency of a thermoelectric generator (Eq. 1.2.1) and the maximum coefficient of performance of a thermoelectric refrigerator are both closely tied to the Z of the material. Z is equal to $\frac{\alpha^2\sigma}{\kappa}$, where α is the TE material's Seebeck coefficient, κ is the thermal conductivity, and σ is the electrical conductivity. T is the mean operation temperature of the device. For approximately 40 years, the maximum figure of merit was approximately ZT~1 (for the Bi_2Te_3 material family). Figure 8 shows several important bulk thermoelectric figures of merit as a function of temperature. The word "bulk" is used from now on to describe any material not a thin-film, that is, whose minimum dimension can exceed 100 microns.

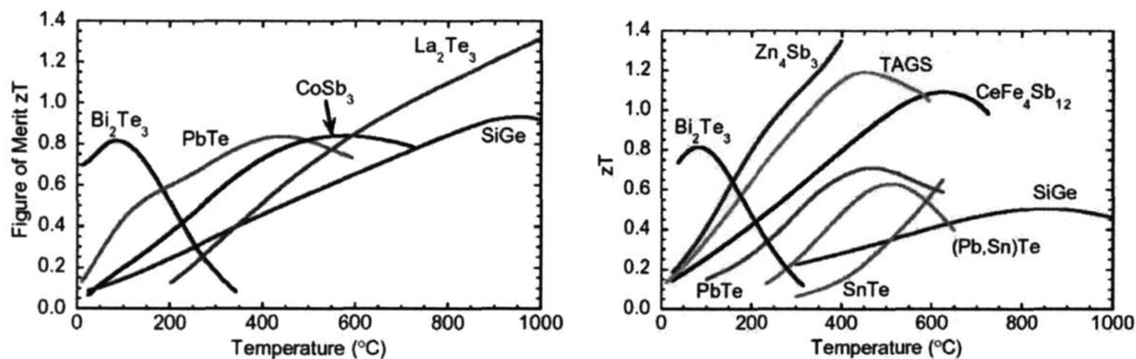


Figure 8. ZT vs. mean operating temperature for several n-type (left) and p-type (right) bulk thermoelectric materials. (Images from Jeff Snyder, http://www.its.caltech.edu/~jsnyder/thermoelectrics/)

There are some systems for which the small efficiencies attainable from bulk thermoelectrics are acceptable. For applications involving waste heat scavenging, the generated power per unit area or per unit cost (rather than the efficiency) becomes the economically important design specification, making it easier for present materials to find practical use[8,9].

One of the largest potential markets for TE power generation is waste heat reclamation in automobiles. Around two thirds of the power from combustion in a typical automobile is lost as heat rejected through the radiator and exhaust. If some of this heat could be used to supplement the electricity demand of the car, overall fuel efficiency increases on the order of a couple of percent may be possible. Several investigations into possible waste-heat reclamation from the exhaust of automobiles and trucks have been conducted. At the present time, prototype automotive TE generators are not able to produce enough power to pay for their cost and justify mass-production, although research and development in this area continues[10]. One

of the major limitations of automotive waste heat reclamation is not just the inherent efficiency of the generator material, but also the system level question of where the waste heat is to be extracted so as to minimally impact current car design, and the low "quality" of the heat at the most suitable heat extraction points. These factors combine to limit the amount of power generated per kilogram of extra weight added to the car, arguably the key performance metric for this application[11]. For example, the exhaust system after the catalytic converter is a relatively easy place to integrate a generator (Figure 9) but the exhaust temperature is lowest there and the inherently low convective heat transfer coefficient from the gas to the hot side of the generator (to say nothing of the cold side) is also a drawback. Extracting heat before the catalytic converter would allow for greater exhaust gas temperatures and higher gas pressures but would adversely affect the operation of the catalytic converter whose efficiency is dependent on a high temperature. For diesel powered cars and trucks the turbocharger typically present before the catalytic converter might also interact adversely with a TE generator. Extracting heat directly from the engine or radiator fluid is an interesting possibility since the metal cylinder block or the radiator fluid both offer very high quality heat sources, but to achieve high enough hot side temperatures the TE generator should probably be integrated in series with the engine cooling circuit very close to or inside of the cylinder block. This would entail a major redesign of the automobile engine cooling system[12] for what would most likely result in only an incremental improvement in overall fuel efficiency. The weight of the generator and its corresponding cost in fuel efficiency must also be considered. The example of automotive TE power generation highlights the importance of considering the total impact of the technology on the cost and operation of the system in which it is embedded.

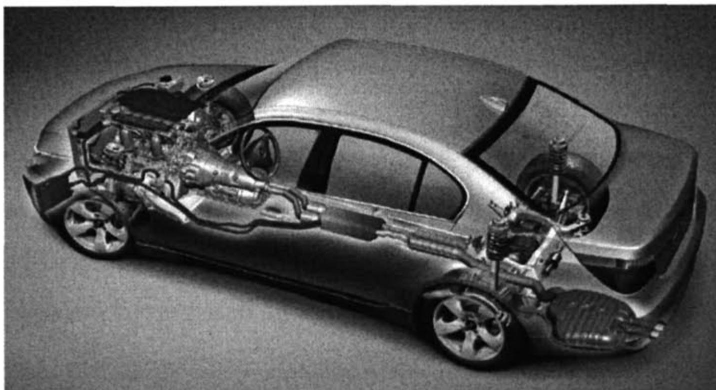


Figure 9. Thermoelectric generator and heat exchanger envisioned for the BMW 5-series (from http://www.greencarcongress.com/2006/09/bmw_introduces_.html).

Adding a TE generator thermally downstream from a car's engine (e.g. on the exhaust pipe) might ultimately improve the car's energy efficiency . Placing a TE generator thermally upstream from another part of a system that requires heat can also improve system efficiency, since the waste heat from the cold side of the TE generator, always thermodynamically, can then be used profitably[13]. An interesting example of a self-contained combustion generator using this idea has been proposed[14]. Here the heat for the generator is produced from a hydrocarbon burner and passed through a thermoelectric generator, producing useful electricity. The heat rejected from the cold side then preheats the input gas for the combustion chamber. Assuming only a relatively modest ZT=1, the resulting total system efficiencies are calculated to be as high as 23%, far higher than those attainable even for idealized conditions using a traditional thermoelectric generator. A MEMS prototype using this idea has been demonstrated[15], although high efficiencies were not achieved.

When thinking about thermoelectric applications, an interesting question is how high the material ZT must be in order to become useful for large-scale power generation and refrigeration applications. A recent estimate by Richman and Stringer[16] tries to compare thermoelectric cooling with conventional vapor compression cycle refrigeration technology using the coefficient of performance (COP), and compares thermoelectric generation with conventional large scale power generation according to the thermal-to-electrical conversion efficiency (given in Eq. 1.1.1 above for thermoelectrics). The COP of a refrigerator is the ratio of cooling power to the input electrical power, at the input electrical power that maximizes the ratio. The thermodynamic maximum COP is the Carnot COP, given by $COP_{Carnot} = \frac{T_C}{T_H - T_C}$, where T_C is the desired temperature inside the refrigerator and T_H is the ambient temperature. Richman and Stringer note that a household refrigerator manages to achieve a COP of 46% of Carnot, and the same performance from a thermoelectric would require ZT~8. A similar comparison can be made for power generation. The total thermal efficiency of a state of the art 400 MW gas turbine combined cycle power plant can exceed 60% [17], and a portable 10 kW gas generator has a thermal efficiency of around 25% [18]. For an assumed temperature difference across the TE generator given by T_H=2000K and T_C=400K, to attain the performance of the small gas generator requires a ZT of around 12, and to attain the performance of the power plant requires a ZT of around 400. Based on this simple analysis, it appears likely that in the near future, the commercial applicability of thermoelectric generators and coolers will be limited to applications that are small enough in size so that conventional technologies cannot compete,

or for which efficiency and COP are not as important. Two examples of such applications are the spot cooling and temperature control of electronic or optoelectronic devices on semiconductor chips or waste heat scavenging for powering small portable electronic equipment.

Recently, a new bulk material (cubic $AgPb_mSbTe_{2+m}$) has been reported with a record figure of merit ZT=2.2 [19]. However, the maximum possible figure of merit from any future bulk material has been theoretically estimated to be around ZT=4 [20]. This limit and recent successes in growing large ZT thin-films suggest that we should look to nanostructured and thin-film materials for major future advances in thermoelectric performance. Thin-film materials have also claimed values of ZT greater than 2 for certain material systems[21,22] and there appear to be few fundamental barriers to greater values of ZT in the future. The semiconductor manufacturing techniques now available for tailoring the properties of materials on the nanoscale may offer the possibility of dramatic improvements in generator efficiency and power density[23]. Figure 10 from a recent summary by Majumdar[24] illustrates the recent dramatic advances in thin-film semiconductor ZT.

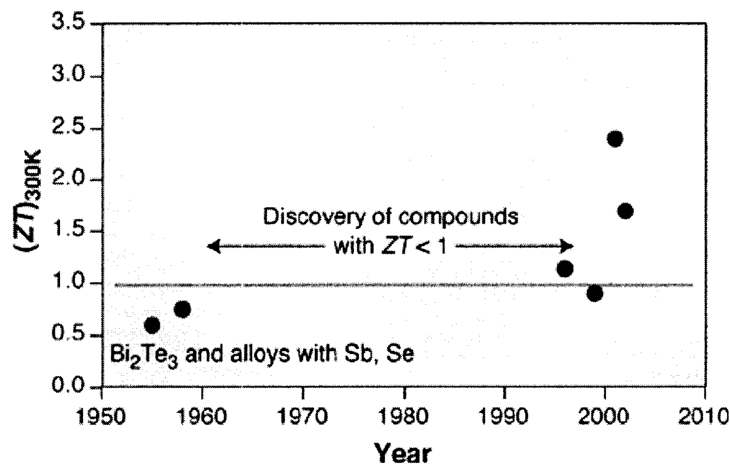


Figure 10. Recent advances in thin-film thermoelectrics. (from Majumdar, 2004)

Shown in Figure 11 is an example of a thin-film thermoelectric material, a lattice-matched doped superlattice consisting of alternating 10 nm periods of Si and SiGeC. The material is designed to reduce of thermal conductivity of the material by increasing phonon scattering at the superlattice interfaces, thereby increasing ZT.

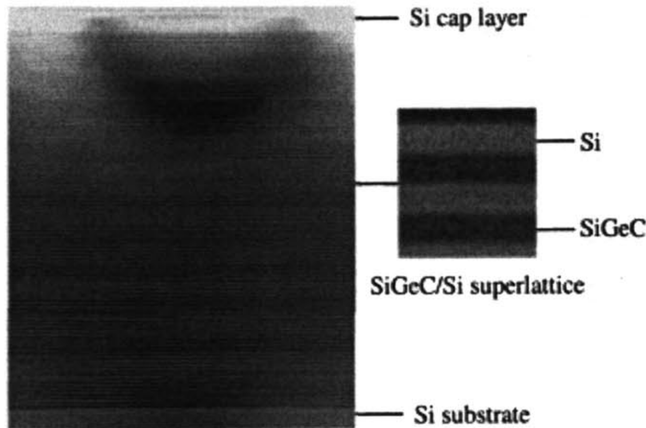


Figure 11. SiGe superlattice thermoelectric material. (from Fan et. al., 2001)

In light of the recent reports of high figure of merit (ZT>1.5) thin films in the scientific literature, some important questions must be addressed. First, how is the thermoelectric figure of merit to be determined for a thin film? Second, how is the thin film to be integrated into a useful generator?

There are two main techniques employed for measurement of bulk TE materials: the transient Harman technique, and the DC ZT-meter method. Both methods demand knowledge of the thermal environment very close the edge of the thin film. Roughly speaking, the Harman technique requires that the thermal resistance to and from both sides of the thin film be known, and the DC ZT-meter method requires that the temperature at both the hot and cold sides of the thin film be known. Unfortunately accurate thermal measurements on the microscale and nanoscale are notoriously difficult. The constraint that the measurement must not perturb the thermal environment motivates us to search for a new method of non-contact thermal characterization suitable for microscale and nanoscale investigations of heat transport in semiconductor materials. Shown in Figure 12 is a chip-level thermal image taken using a commercially available IR-camera.

While allowing non-contact thermal measurement, the thermal resolution of infrared imaging systems is typically around 3-5 μm due to the thermal emission spectrum of typical objects near room temperature, the availability of detectors in the desired spectral range, and the fundamental constraints placed on infrared far-field imaging due to the diffraction limit of light. A new non-contact thermal measurement technique applicable in the micro and nano regime is needed to study the heat transfer through typical thin-film elements.

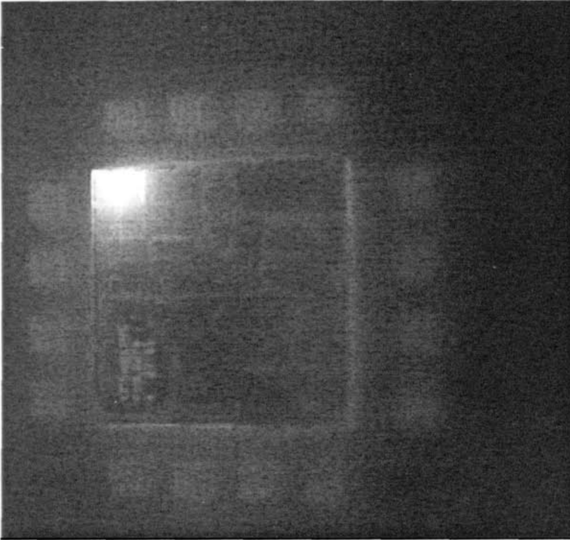


Figure 12. Infrared image of a computer chip fault during real conditions testing using the FLIR Systems ThermaCAM (http://www.flirthermography.com/success/ir_image/1030/industry_id/1000/).

One promising candidate technique for non-contact thermal measurement is thermoreflectance imaging [25,26]. The technique relies on the fact that in most solid materials, the optical index of refraction is a function of temperature. For small temperatures changes the normalized change in reflectance at the interface between air and the material is then proportional to the change in temperature at the interface. If the constant of proportionality is measured for the material(s) of interest, a spatial image of the temperature may be constructed from a corresponding image of the normalized reflectance. The technique has important advantages over IR imaging. The spatial resolution of the measurement is limited by the diffraction limit of reflected visible light rather than the diffraction limit of emitted infrared thermal radiation. Together with the availability of small pixel pitch CCD imaging chips and high a numerical aperture optical microscope train, this pushes the potential spatial resolution into the submicron/nanoscale regime. Thermoreflectance also works well on metallic or highly reflective materials that cause difficulty for IR imagers due to their correspondingly low emissivities. A potential drawback of thermoreflectance imaging is that it is a modulation technique so that it measures the change in reflectivity with temperature rather than absolute temperature directly. The typical fractional change in reflectivity per degree Kelvin of temperature is also rather small (~1/10000), which might suggest that in spite of the excellent spatial resolution the measurement might never attain the kind of thermal resolution available from traditional lock-in IR imaging (~10s of mK). We will show otherwise later in this thesis.

Once the thin film has been characterized, it remains to demonstrate that it can be usefully employed for power generation. The most glaring issue to be addressed is that an excellent TE material might have a third the thermal conductivity of a conventional bulk TE material, but due to constraints on the maximum practical growth thickness, it might be only possible to grow it to a small fraction (e.g. $1/60^{th}$) of the thickness of the standard bulk material. This complicates generator design since per unit area (in the cross-plane direction) the elements will conduct 20 times more heat for a given temperature difference. For a generator to function properly heat must be delivered to the hot side and removed from the cold side fast enough to maintain the desired temperature gradient which in turn determines efficiency and power. Designing a generator with the thin film elements demands either a radically different packaging approach to force the elements into more traditional aspect ratios (e.g. multiple stacked cross-plane elements or elements rotated somehow such that heat transfer is in the longer in-plane direction), or greatly enhanced heat transfer mechanisms on the hot and cold side of the generator.

1.3 Thesis Outline

This thesis focuses on the implications and use of new high ZT materials for thermoelectric power generation and on the development of a new measurement tool (thermoreflectance) used to evaluate their properties and performance.

In Chapter 2 we present an element level study on the use of thin-film high-Z elements for power generation. TE generation is compared with other types of heat engines and the case for optimizing power density rather than efficiency is made. The constraint imposed by the heat sink (and finite heat transfer in general) are discussed and it is found that changes to the standard theory of optimized thermoelectrics are required in order to optimize generators given present materials and heat-sink technologies. Directions for future system improvement are discussed here in light of the theory and it is found that the figure of merit ZT is not always the best way to understand the effects of future material improvements in thin-films on power generation.

Having identified heat removal from the cold side of the generator as a limiting factor to higher generator efficiency and power density Chapter 3 examines a few candidate thermal management techniques in detail along with their effects on system performance. Heat

spreading is examined as a means of trading off efficiency and power density and as a way of dealing with the large power densities resulting from thin-film elements. The experimental characterization of a heat sink is demonstrated. The feasibility of microchannel cooling at the cold side of the generator is theoretically investigated, including an analysis of the total system efficiency and power budget. Finally, a design of a generator system is developed using the models developed for each of the system components, and we show that substantial improvements in power density are possible over traditional power generation systems.

In Chapter 4, thermoelectric metrology is discussed and a candidate method for TE characterization (the ZT-meter) is proposed and demonstrated. Single elements made from state of the art thin-film materials are tested under large applied temperature differences for both generated power density and efficiency. To test how well system level performance could be extrapolated from measurements on individual TE elements, a commercial generator was disassembled, its n and p elements were separately tested in a custom built ZT-meter apparatus, and the predicted packaged generator performance was compared with the actual generator performance. A range of commercial and research grade thermoelectric materials were then tested in the apparatus. General issues and limitations relating to ZT-meter measurements (parasitic thermal interface resistances, electrical contact parasitics) are discussed in detail, as are issues relating to other standard measurements on thermoelectrics (transient Harman, Van der Pauw).

In Chapter 5, the thermoreflectance imaging technique is developed, characterized, and modeled. A detailed discussion of issues related to accuracy and calibration is given. The technique is used to thermally image the operation of thermoelectric materials and other devices with sub-micron spatial resolution and milli-Kelvin temperature resolution. With an associated time-domain thermoelectric simulation tool, the method also provides a means for non-contact thermal measurements on thermoelectric elements.

In Chapter 6, we discuss some general conclusions and propose a few questions to motivate directions for future research on thermoelectric power generation and thermoreflectance imaging.

Chapter 2: Theory and optimization of high-power density thermoelectric elements

2.1 Overview

In Chapter 2 we present an element level study on the optimal use of thin-film high-Z elements for power generation. Chapter 2.2 presents the basic expressions for TE power generation that are found in the literature. Chapter 2.3 compares TE power generation with a more common heat engine (the Brayton cycle), and discusses the fundamental nature of the trade-off between generated power and efficiency in a real engine. We show why optimum power is often the more rational choice for design. Chapter 2.4 examines the constraint imposed on thermoelectric power generation by the heat sink (and more generally, finite rate heat transfer), and Chapter 2.5 incorporates the heat sink into the standard model. Chapter 2.6 discusses the finite difference solver that was written to simulate TE element performance in this thesis. In Chapter 2.7, the solver is applied to the problem of selecting a load resistance that optimizes power generation, and we find that the optimum load when a heat sink is included in the device model can be quite different from that which is given by the standard theory. Chapter 2.8 and 2.9 perform similar optimizations for the element length and material properties, respectively. It is found that for a given heat sink technology, there is an element length that maximizes power generation whose value can be approximated by appealing to a thermal impedance matching criteria similar to the familiar electrical impedance matching criteria. It is also found that when the heat sink is important, the figure of merit ZT is not the best way to understand the effects of future TE material improvements on power generation.

2.2 Basic expressions for TEG

The designer of a TE generator has a choice between optimizing the efficiency or the power density of the device. Material improvements (increases in Z) will typically improve both metrics, while other choices (geometry, heat sinking, and load) can allow efficiency to be traded off for power density or vice versa, depending on the application. In a portable power application for which a generator is operated from a small gas burner one might be most concerned with the efficiency of the generator, which limits the total fuel efficiency of the power source. On the other hand generating a large power density is more important in applications

involving waste heat recovery. In this case, since the power running the generator will be wasted anyway one is simply concerned with generating as much power as possible from the available heat source. To understand these issues, it's useful to review the most basic expressions for the generated power and the efficiency of a thermoelectric generator (TEG)[27,28].

Consider heat and current flow in the device (Figure 13). The TEG is a heat engine relying on a difference in temperature between hot and cold reservoirs to drive electronic transport through an external load. We assume a hot side temperature of T_H and a cold side temperature of T_C. T_F denotes the temperature of the cold thermal reservoir, typically an ambient fluid or gas. The heat sink keeps T_C as close to T_F as possible. One could choose to allow T_H to differ from the temperature of the hot reservoir, but this will not be done here since modeling a heat concentrator is similar to modeling a heat sink. Also for simplicity, only a single TE couple (leg) of area A and length L is optimized here. The optimization method used here can be extended to handle a p-doped and an n-doped leg together.

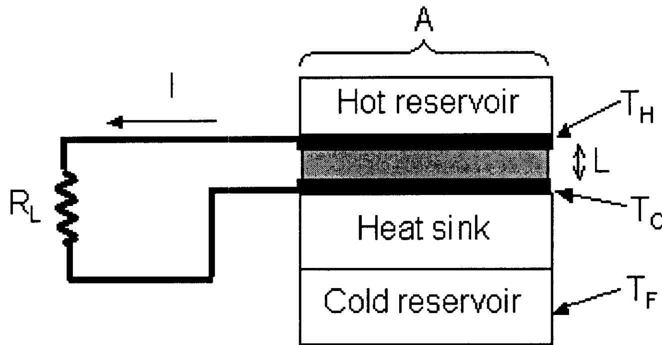


Figure 13. Schematic diagram of a thermoelectric generator.

The heat entering the TEG from the hot side is given by the sum of the conduction power through the device, the Peltier heat power carried by the electrical current at the hot side, and the Ohmic heat rejected from the device.

$$P_{in} = P_{cond} + P_{Pelt} - P_{Ohmic}$$

$$P_{in} = \frac{\kappa(T_H - T_C)A}{L} + \alpha T_H I - \frac{1}{2} I^2 R_{int} \qquad (2.2.1)$$

where $I = \dfrac{(T_H - T_C)\alpha}{R_{tot}}$. $\qquad$ (2.2.2)

The expression for the current I follows from the fact that the voltage across the TEG will be the Seebeck coefficient of the TE material (α) times the temperature difference across the element. The expression for the resistances R_{int} and R_{tot} can be written in terms of the TE material's conductivity σ, the chosen load resistance R_L, and the geometry of the element.

$$R_{tot} = R_{int} + R_L = \frac{L}{\sigma A} + R_L . \qquad (2.2.3)$$

The generated power density is simple I^2R_L/A. The load maximizing the generated power density is matched to the internal resistance (R_L=R_{int}). This results in an expression for power density

$$Q_{gen,P} = \frac{(T_H - T_C)^2 \alpha^2}{4L/\sigma} . \qquad (2.2.4)$$

Note that the Seebeck coefficient α and the electrical conductivity σ show up directly in the expression for generated power density. Decreasing the device length also increases the generated power density. The subscript P reminds us that a choice was made to optimize power density rather than efficiency. Since the generated power density and the input power density are both known, the efficiency can be computed. After some algebraic simplification, we find

$$\eta_P = \frac{P_{gen}}{P_{in}} = \frac{T_H - T_C}{2T_H - \frac{T_H - T_C}{2} + \frac{4}{Z}} = \eta_C \frac{1}{2 - \eta_C/2 + 4/(ZT_H)} . \qquad (2.2.5)$$

The efficiency depends on all three basic material parameters (α, σ, κ) through the lumped parameter Z, as well as the hot side temperature and the Carnot efficiency η_C. Increasing Z increases the efficiency up to the fundamental limit set by (but less than) η_C.

Rather than picking a load to optimize generated power density we can also choose a load to optimize efficiency. The resulting optimum load differs from the load for maximum power generation by a factor μ:

$$R_L = \mu R_{int} \equiv \left(1 + Z\left(\frac{T_H + T_C}{2}\right)\right)^{1/2} R_{int} . \qquad (2.2.6)$$

This leads to an expression for power density

$$Q_{gen,\eta} = \frac{(T_H - T_C)^2 \alpha^2}{L/\sigma} \frac{\mu}{(1+\mu)^2} \tag{2.2.7}$$

and an expression for an efficiency

$$\eta_\eta = \frac{P_{gen}}{P_{in}} = \frac{T_H - T_C}{T_H} \left[\frac{\mu - 1}{\mu + \frac{T_C}{T_H}} \right]. \tag{2.2.8}$$

Note that in the limit of large ZT the bracketed term in the expression for efficiency approaches unity. What remains is simply the Carnot efficiency of an ideal heat engine. To get a sense for the difference between the two choices for optimum load note that for ZT=1.5 (currently achievable) the factor μ relating the loads is 1.58. This is big enough to be of practical importance in device design and will increase as thermoelectric materials continue to improve.

So far in this discussion we tacitly neglect thermal and electrical contact resistances and assume a perfect heat sink ($T_C=T_F$). In other words, the temperatures of the hot and cold sides remain fixed and do not depend upon device operation.

2.3 Optimum power or efficiency?

The trade-off between optimal power and optimal efficiency can be clarified using a type of plot often used to express the effects of internal and external irreversibilities on thermodynamic cycles. For a given temperature difference, a heat engine can typically be operated at a variety of powers and efficiencies. By varying an operational parameter, a curve of the efficiency vs. power is traced out. To better understand thermoelectric generators in the context of general heat engines we first consider a traditional engine.

The plot in Figure 14 shows a family of efficiency-power curves for the Brayton cycle[29], the thermodynamic cycle associated with gas turbine engines. Power here is normalized, and denoted by P/P^*_{max}. The Carnot efficiency for a perfectly efficient cycle has been set at 80%.

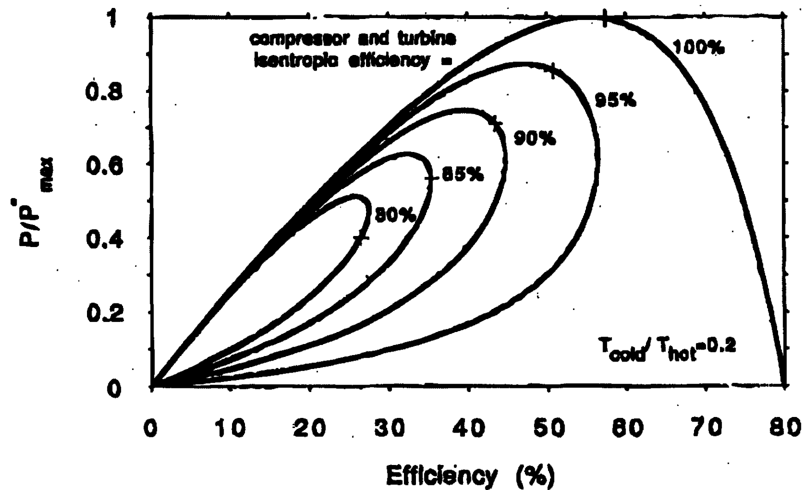


Figure 14. Power vs. efficiency curves for gas turbine engine with varying internal frictional losses, demonstrating general features of heat engines with internal irreversibilities.

The Brayton cycle works using a compressor, a reactor, and a turbine. Air is drawn in through the compressor, mixed with fuel and burned in the reactor, and the resulting expanding exhaust gas turns a turbine. The power from the turbine is used to run the compressor and also produce useful excess (e.g. electrical) power. The inefficiencies (here regarded as internal losses due to fluid friction) in the compressor and turbine can be described using a parameter called the isentropic efficiency. Each of the closed curves in Figure 14 corresponds to a different isentropic efficiency varying between 80% and 100% in steps of 5%. For a given isentropic efficiency the efficiency-power curves are constructed by varying the pressure ratio in the compression and expansion branches of the cycle and plotting the efficiency and power at each operation point. Most of the curves form loops. However, the two ends of the curve terminating at the origin actually correspond to different physical situations. In the limit of small pressure differences between the inlet (or outlet) gas and the hot reactant gas very little power can be generated from the turbine since there is no pressure to drive its motion. But the heat leak through the system is finite since the fuel is still being burned at some finite rate, making the efficiency approach zero as well. This limit corresponds to the part of the curves terminating at the origin from the counter-clockwise direction.

In the limit of high pressure ratio the efficiency-power of curves of Figure 14 approach the origin from the clockwise direction. The one exception is the 100% isentropic (lossless) curve which terminates at the zero power but Carnot efficient point. The intuitive picture is that one is

supplying a large power into the compressor and hoping to recover this from a correspondingly large power out of the turbine. This strategy pays off only if the efficiencies of the turbine and compressor are nearly perfect, since there is only so much energy available from the burning fuel to run the cycle. Even in that ideal case one pays a price in that the large stroke volumes required correspond to very long engine cycles (slow engine speed) so that very little power is produced in this limit.

Analogous curves can be traced out for a TE generator by assuming fixed temperatures and thermoelectric material properties while varying the load resistance (Figure 15).

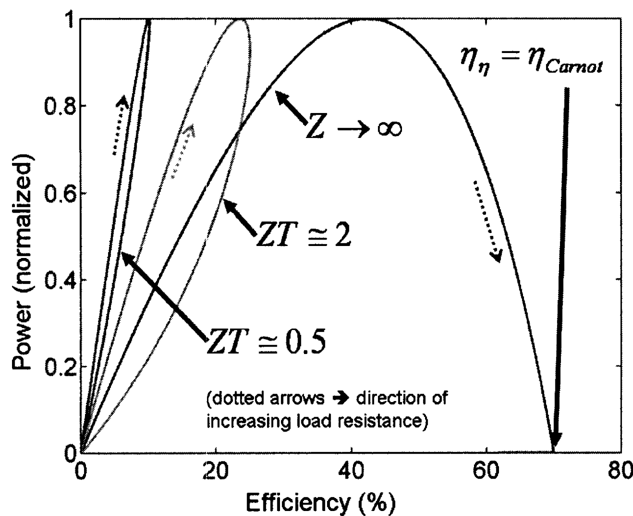


Figure 15. Normalized power vs. efficiency curves for TE elements with various figures of merit. The rightmost curve allows Z to approach infinity, corresponding to a perfect material with no internal losses. The curves are generated by varying the external load resistance. Note that for typical Z's, the optimum efficiency and optimum power points are close to one another.

Here the load resistance plays the part of the pressure differential in the Brayton cycle. With no voltage across the external load of the TE element there can be no work done, just there is no work done by a turbine which sees no pressure differential. The finite thermal conductivity of the TE element plays the role of the finite heat leak of the Brayton cycle, driving the efficiency of the TE element to zero in the short circuit limit. Likewise, the situation analogous to a high pressure differential limit of the Brayton cycle is the open circuit limit of a thermoelectric generator at which no power is produced, and the efficiency (assuming finite thermal conductivity) is zero.

According to Eq. 2.2.6 for an ideal TE material (high ZT limit) the optimum efficiency load is larger than the optimum power load. The situation is similar in the Brayton cycle where the pressure differential yielding optimum efficiency is higher than the pressure differential yielding maximum power. For smaller internal efficiencies the maximum power and efficiency points are located closer to one another. It is clear from comparing the ZT=1 and ZT=2 curves to the ideal ZT=infinite curve in Figure 15 that modern thermoelectric materials are still in the low isentropic inefficiency limit relative to the Brayton cycle. We also note that in the case where perfect internal efficiency was assumed, in both engines there is a limit corresponding to very low power output at which the efficiency approaches the Carnot value. Even ideal engines only achieve the Carnot efficiency in the limit of infinitesimally small power generation, making the Carnot efficiency (or "fraction of Carnot" efficiency) a somewhat odd metric for understanding the ultimate limits of practical power generation.

Recent theoretical work in linear irreversible thermodynamics[30] has established that the efficiency at maximum power of a perfectly internally reversible (Carnot efficient) engine whose heat transfer to thermal reservoirs is constrained by the Onsager relations is equal to $1-\sqrt{T_C/T_H}$, as compared with the Carnot efficiency of $1-T_C/T_H$. This is the so-called Curzon-Ahlborn limit. We suggest that the efficiency at maximum power (as opposed to at maximum efficiency) is the more rational limit to consider when discussing the best conceivable operating point of a generator, since at the maximum power point both efficiency and power can be nonzero [17,31]. More generally, we suggest that the optimum operation point of a perfect generator would be found somewhere between the Curzon-Ahlborn limit on the low end and Carnot efficiency on the high end [17]. To get a better estimate on the ideal operating point beyond that bound one must presumably consider other (e.g. economic) concerns. For example, a high cost of fuel relative to other costs could push the rational operation point towards maximum efficiency, whereas a free source of fuel (e.g. geothermal or scavenged waste heat) might put the optimum closer to maximum power. To optimize the practical (as opposed to ideal) performance of a thermoelectric generator, one must consider the other physical constraints specific to TE generation, including finite heat-leaks, finite heat transfer between reservoirs, and finite internal irreversibilities. We turn to that task next.

2.4 High Z thin-film TE materials – the heat sink limited regime[32]

High Z thin-films offer the prospect of substantially larger generated power densities provided that some significant obstacles are overcome[33]. Epitaxial growth of thick TE material (>> 10 μm) is difficult and expensive. The deleterious effects of parasitic electrical and thermal resistances are more severe for short devices[28,34]. With large temperature gradients, issues related to the non-uniform thermal expansion of the TE materials, electrical contacts, and packaging become more severe. There is also the daunting problem of connecting many TE couples electrically in series and thermally in parallel in order to obtain useful voltages, and the requirement for large heat powers to be removed at the cold side of the generator[35]. Along with these other issues, we have found that the most fundamental problem for thin-films is the large heat power that must be rejected at the cold side through the heat sink. We argue that at the present time, heat sinks fully compatible with the new thin-film materials do not exist. We therefore focus on the challenges associated with designing thin-film TE generators given this constraint. We will look at optimizing the power (as opposed to the efficiency) for the reasons discussed in the previous chapter.

Heat sinks	Heat transfer coefficient h [W/cm^2/K]	Thermal resistance R_{th} [K/W] (1 cm^2 area)	TE elements/generators	Thermal resistance R_{th} [K/W] (1 cm^2 area)
natural convection (air)	0.0005	2000	PbSe-PbTe (K_{tot}=0.0062 W/cm/K, L=104 μm)	1.68
natural convection (water)	0.01	100	Bi_2Te_3-Sb_2Te_3 (K_L=0.0022 W/cm/K, K_e=0.0053 W/cm/K, L=5.4 μm)	0.0720
forced convection (air)	0.004	250	SiGe-SiGeC (K_{tot}=0.085 W/cm/K, L=2 μm)	0.00235
forced convection (water)	0.6	1.66	Hi-Z HZ-2 (bulk) commercial generator (K_{tot}=0.024 W/cm/K, L=0.508 cm)	21.2
pool boiling (water)	2	0.5		
microchannel heat sink	20	0.05		

Table 1. Thermal resistances of various heat sinks and thermoelectric materials.

It has long been understood that a device made with thin-films will be limited as much by heat removal as by intrinsic material quality[27]. Table 1 compares thermal resistances across different heat sinks[36] with those across some representative state-of-the-art thin-film TE elements[21,22,37,38]. Note that the thermal resistance R_{th} of a heat sink is given by $R_{th} = \frac{1}{hA}$ where h is the heat transfer coefficient and A is the area of the heat sink.

The table allows for the comparison of the temperature drops across the TE legs and heat sinks if they are connected in series between a hot and cold reservoir as shown in Figure 13. As will be shown later, for a fixed heat sink, near the optimal operation point for maximum power density the temperature drop across the TE leg and the heat sink are roughly equal. Note that two of the references in Table 1 are for materials with leg lengths under 10 μm[22,37]. The other is unique in that is has been grown to lengths over 100 μm[21]. It is clear from Table 1 that at the present time, only the most aggressive heat sink technologies can handle the large heat fluxes required. For example, both the Bi_2Te_3-Sb_2Te_3 and the SiGe-SiGeC have such small thermal resistances that only the microchannel heat sink is capable of matching their heat removal requirements. While promising in terms of cooling performance, microchannel heat sinks have never been demonstrated for this application and the power consumed by actively pumping them (or for any of the forced convection heat sinks) is a major concern for power generation applications. Table 1 shows that given the current limitations on thin-film growth some of these new materials will be forced to operate in a heat sink limited regime, meaning that the temperature drop across the heat sink will be much greater than that across the thin-film. While the challenges of working with such high heat power densities are obvious, it should be recalled from Eq. 2.2.4 that if the power from the thin generators can be handled (i.e. the cold side temperature can be kept from becoming too hot) we are rewarded with greater generated power densities. To this end, we reexamine the standard conception of optimized thermoelectric generation to understand how performance is altered in this heat sink limited regime.

2.5 Improved model for thin-film TE generator

It is well known that the standard model described in Chapter 2.2 neglects several important parasitics. Between the metal leads on the TE material and the material itself (typically a doped semiconductor or semimetal) is an Ohmic contact. For thick devices, the contact resistance R_C (ranging between 10^{-4} and 10^{-8} Ω cm^2) can sometimes be neglected. For a thin-film device this is not the case. The contact resistance can be a significant fraction of or even exceed the

internal resistance. This is a particularly difficult issue for contacts capable of surviving the high absolute temperatures large thermal expansions inherent in TE generator operation. With regards to its effect on device performance R_C is effectively added to the load resistance, but the power lost in the contact is wasted heat.

The total series electrical resistance in the TE circuit is given by the sum of the internal and load resistances,

$$R_{tot} = R_{int} + R_L = \frac{L}{\sigma A} + \frac{2R_C}{A} + R_L . \qquad (2.5.1)$$

Here A is the area of the TE generator. The factor of two is because we assume two contacts per element.

The effects of the heat sink and any thermal resistance on the cold side are modeled by a lumped heat transfer coefficient h so that the total power transferred through the heat sink is $hA(T_C\text{-}T_F)$. It is assumed here that the area of the heat sink is that of the device. Although it is possible to trade off power density for heat sink performance with a heat spreader, in many cases the challenges of packaging such a system outweigh its benefits and for simplicity we begin our analysis by ignoring heat spreading. (In Chapter 3 we will assess the possible benefits of heat spreading). We also neglect any hot side thermal resistance since it would be modeled in an identical manner to the heat sink on the cold side.

Included in the electrical and thermal contact resistances are the effects of any interconnects needed to wire the elements together. In a real generator, hundreds or thousands of elements are wired together electrically in series and thermally in parallel, introducing additional parasitics (Figure 16). These package parasitics can simply be added to the contact resistance and heat sink coefficient in the model. Once both an n-type and a p-type element are simulated, the performance of the entire generator can be obtained by scaling the power generated in a single p-n couple by the number of couples in the full package.

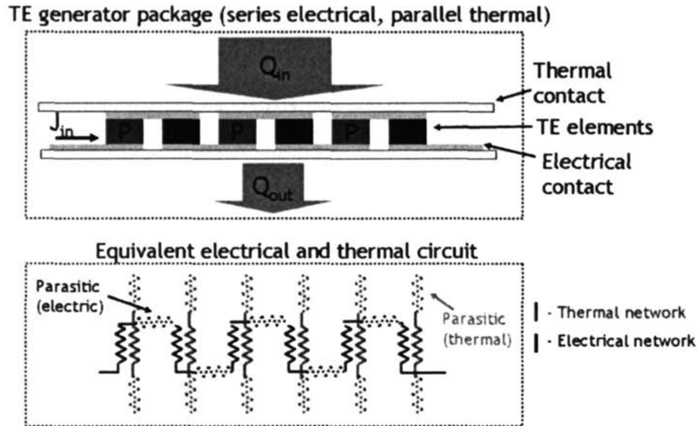


Figure 16. Additional parasitics due to packaging that can be easily included in the model using an effective contact resistance and an effective heat sink parameter.

Proceeding along the same lines as Chapter 2.2, the power density generated by the TE element is now given by:

$$\eta = \frac{Q_{gen}A}{\frac{\kappa(T_H - T_C)A}{L} + \alpha T_H I - \frac{1}{2}I^2R_{int}} \tag{2.5.2}$$

$$Q_{gen} = \frac{(T_H - T_C)^2 \alpha^2}{R_{tot}^{\ 2}} \frac{R_L}{A}. \tag{2.5.3}$$

These equations are complicated by the dependence of the cold side temperature T_C on the operating point of the generator. This dependence arises because changes in R_L alter the amount of current flowing through the generator. This changes both the Ohmic heat and the Peltier heat rejected into the heat sink, changing T_C. The cold side temperature of the generator has become a bias-dependent quantity.

Equating the heat leaving the cold side of the TE leg with the power entering the heat sink, we obtain

$$\begin{array}{ccccccc} P_{cond} & + & P_{Pelt} & + & P_{Ohmic} & = & P_{sink} \\ \frac{\kappa(T_H - T_C)A}{L} & + & \alpha T_C I & + & \frac{1}{2}I^2R_{int} & = & (T_C - T_F)hA \end{array} \tag{2.5.4}$$

$$\text{where } I = \frac{(T_H - T_C)\alpha}{R_{tot}}. \tag{2.5.5}$$

Eq. 2.5.1 and Eq. 2.5.5 can be substituted into Eq. 2.5.4 to solve for T_C. Using the resulting T_C in Eq. 2.5.2 and Eq. 2.5.3 relates the efficiency and generated power density to material parameters, the length, the heat sink (h), and the load R_L. These relations, too complex to bother writing in closed form here, allow the effects of different system design choices on performance to be studied.

In what follows, we choose to study 100 μm of n-type SiGe at 900K[39]. This corresponds to α=247 μV/K, κ=0.0393 W/cm/K, and σ=485 1/ Ω/cm. The area of the couple is 1 cm^2. The heat sink parameter is h=0.5 $W/cm^2/K$, consistent with forced water convection[36]. Also included is an electrical contact resistance of 10^{-6} Ωcm^2. The temperatures are T_H=923 K, and T_F=300 K.

2.6 Finite difference solver for Onsager relations

The equations describing the flow of energy and charge (electricity) in a standard thermoelectric material are the Onsager relations, the basic equations of first-order non-equilibrium thermodynamics[40]. The equations can be written in a form amenable to simulation:

$$\begin{aligned} \vec{J} &= \sigma\vec{E} - \sigma\alpha\vec{\nabla}T \\ \vec{Q} &= \alpha\vec{J}T - \kappa\vec{\nabla}T \end{aligned} \quad (2.6.1)$$

Here J is the particle flux (electrical current) and Q is the energy flux. E is the electric field resulting from any gradient in the electrical or electrochemical potential. T denotes the temperature distribution in the solid, and σ, κ, and α are the usual thermoelectric parameters. There are also two additional constraining equations resulting from energy and particle number conservation:

$$\begin{aligned} \vec{\nabla}\cdot\vec{J} &= 0 \\ \vec{\nabla}\cdot\vec{Q} &= \vec{J}\cdot\vec{E} \end{aligned} \quad (2.6.2)$$

The equations 2.6.1 are written here for isotropic materials, although they can be generalized using tensor transport coefficients. They are also valid only when the system at hand can be considered in local thermodynamic equilibrium. In the language of statistical mechanics this means that the electron distribution function can be written in terms of the equilibrium Fermi distribution plus a small deviation term representing the departure from equilibrium. The use of the equations (and of the associated thermoelectric parameters σ, κ, and α) is somewhat suspect for processes that are inherently far from equilibrium, e.g. thermionic emission. However, for nonlinear transport it is sometimes the case that the transport can be linearized

and "effective" values of σ, κ, and α can be found[41,42]. In this case Eq. 2.6.1 and 2.6.2 can still be used with the effective parameters. For the thin-films of interest to us (and for typical thermoelectric generator designs) transport is essentially one dimensional allowing the transport to be represented with three coupled first-order differential equations with non-constant coefficients:

$$J = \sigma(x,T)E(x) - \sigma(x,T)\alpha(x,T)\frac{dT}{dx}$$
$$Q(x) = \alpha(x,T)JT(x) - \kappa(x,T)\frac{dT}{dx} \qquad (2.6.3)$$
$$\frac{dQ}{dx} = JE(x)$$

These equations (along with the necessary boundary conditions) contain the physics needed to explain 1-D thermal and electrical transport for arbitrarily graded materials, with arbitrary material parameter temperature dependence. The standard bulk equations (2.2.1. and 2.2.2) can be shown to be solutions of these equations for the case of uniform material parameters and zero Thompson effect $\left(\frac{d\alpha}{dx} = 0\right)$. We are unfamiliar with commercial software which addresses the solution of this system of equations with the needed flexibility, and so we wrote our own using first-order finite difference techniques with a Newton method equation solver using MATLAB (Figure 17). The solver calculates the temperature and heat flow on a one dimensional spatial mesh of N points. There have been some similar numerical methods described in the literature[43].

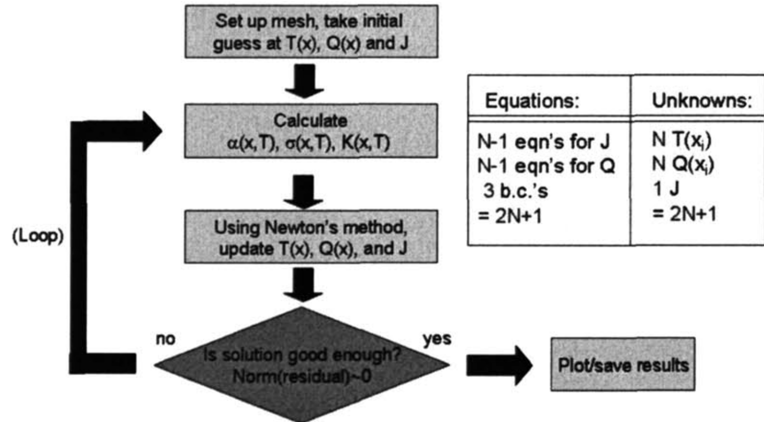


Figure 17. Finite difference algorithm used to compute 1-D transport including arbitrary grading and temperature dependence of material parameters.

The solver allows the performance of functionally graded thermoelectric materials to be examined and optimized. These are materials whose thermoelectric parameters are intentionally varied across the direction of transport in order to enhance performance, typically by varying the doping or the alloy concentration of one of the constituent elements[44]. This is useful for high power thermoelectric generators because it is frequently the case that a thermoelectric material will only perform well over a small temperature range. This is at odds with the requirement that large amounts of power be generated, since generated power should scale roughly with the square of the applied temperature difference. By altering the thermoelectric material parameters across the direction of heat transport in order to re-optimize the material for the local temperature, better performance can often be realized. Recently, a unified formalism has been developed for selecting the right combinations of materials to be used in segmented and cascaded generators[45].

The solver is also useful in examining questions regarding the best methods (and accuracy) for averaging parameters whose material parameters are known to vary with position and/or temperature, so that the resulting effective parameters can be used in the simple bulk equations 2.2.1 and 2.2.2. This is handy for simplifying optimization routines for which repeated solutions using the full finite-difference solver would be cumbersome. It turns out that averaging the Seebeck coefficient, the thermal resistivity, and the electrical conductivity generates effective material parameters that work remarkably well to predict the performance of a spatially graded structure, even for large material spatial dependences. An accurate and universally applicable scheme for handling temperature dependence is more difficult. In most cases of interest (and particularly when the thermal conductivity does not depend too much on temperature or position) one can reasonably assume a linear temperature gradient through the device. One can then use the known temperature dependence of each parameter to generate spatially varying material parameters. The averaging discussed above can then be applied with reasonable accuracy ($< 2\%$ error for the elements we tested).

2.7 Optimum load in heat sink limited regime

For a realistic generator one can use Eq. 2.5.3 in order to find the load resistance R_L maximizing the generated power density Q_{gen}. A useful implicit expression can be written:

$$R_{L,opt} = R_{int} + \frac{2R_L(R+R_L)\frac{d}{dR_L}(\Delta T)}{\Delta T}, \quad (2.7.1)$$

where $\Delta T = T_H - T_C$.

It is clear from this expression that when the temperature drop across the TE element depends on the load resistance (or equivalently the current) then there will be a correction term to the traditional impedance matching equation. This correction arises from the finite heat sink assumed in the realistic model. Because increased current through the device causes the temperature difference ΔT across the TE element to drop, the optimum current is somewhat lower than for the ideal case. Equivalently, the optimum load is somewhat larger. This point has so far been ignored in the literature, likely because for macroscopic (and hence non-heat-sink limited) generators the effect is small.

Before solving for the general case, it is instructive to look at the case of a perfect (internally reversible) thermoelectric generator with an imperfect (also called finite-rate) heat sink. Similar analyses have been performed by others[46]. Referring to Eq. 2.5.4 and 2.5.5 we can rewrite the energy balance equations for a perfect (no internal entropy generation) substance, e.g. $\kappa \to 0, \sigma \to \infty$, as

$$\begin{aligned} \alpha T_C I &= (T_C - T_F)hA \\ P_{gen} &= IV = I\alpha(T_H - T_C) \end{aligned} \quad (2.7.2)$$

The generated power can then be expressed in terms of the reservoir temperatures

$$P_{gen} = IV = I\alpha\left(T_H - T_F\frac{hA}{hA - \alpha I}\right). \quad (2.7.3)$$

When this expression is maximized with respect to the operating current I, one finds:

$$\begin{aligned} I_{opt} &= \left(1 - \sqrt{T_F/T_H}\right)\frac{hA}{\alpha}, \\ P_{opt} &= \left(\sqrt{T_H} - \sqrt{T_F}\right)^2 hA, \\ \eta_{opt} &= 1 - \sqrt{T_F/T_H}, \\ R_{L,opt} &= \frac{\alpha^2 T_H}{hA}. \end{aligned} \quad (2.7.4)$$

This simple case illustrates that the optimum load for power generation is not always impedance matched to the internal resistance of the generator as assumed in the standard treatment.

(Recall that the internal resistance here is zero.) In the case of large Z, the efficiency at maximum power approaches the so-called Curzon-Ahlborn limit of $1-\sqrt{T_F/T_H}$ [31], independent of the heat sink's effectiveness (parametrized here by h). Recall that the Curzon-Ahlborn limit is the efficiency at maximum power of a Carnot engine. Interestingly, this limit is different than (and strictly greater than) that given by setting $Z \to \infty$ in Eq. 2.2.5, equal to $\eta_C \frac{1}{2-\eta_C/2}$, where η_c is the Carnot efficiency. We believe this can be explained by the assumption of impedance matching which is implicit in 2.2.5, which does not necessarily maximize power in the limit of $Z \to \infty$.

To find the optimum R_L for a general material, one can use Eq. 2.7.1 to generate successive corrections to the ideal impedance matched case or one can use a computer to solve the equations exactly. The results of the latter solution are depicted in Figure 18. The lower curve plots the power density versus R_L/R_{int} for the baseline SiGe and the upper curve does the same for a hypothetical material with twice the Seebeck coefficient (but otherwise identical). The curves were obtained using the finite difference solver discussed in Chapter 2.6. The two points marked by an "x" on each curve were calculated using the algebraic optimization procedure described above, and agree quite well with the results of the finite difference simulation. The load for maximum power output is larger than the impedance matched result (dotted vertical line) used in the literature because a lower operating current will lessen the burden on the heat sink. The load optimizing power calculated here also differs from the load maximizing efficiency, which has been addressed elsewhere[28,47]. The top curve in Figure 18 shows that future improvements in Z (in this case, the Seebeck coefficient) will accentuate the difference between the true optimal and the traditional optimal (matched) impedance, with 30% less power being generated in the latter case.

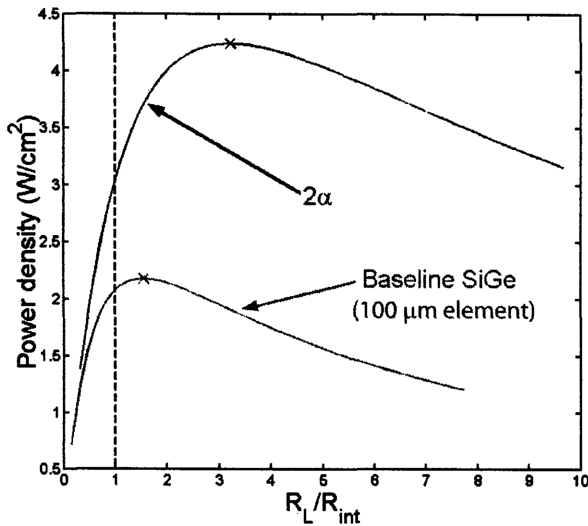


Figure 18. Optimal impedance matching when heat sink limited.

2.8 Optimum length in heat sink limited regime

For a given generator material and heat sink there is a length optimizing the power generated per unit area of the leg. A rough estimate of the optimum length can be obtained by approximating all of the heat transfer in the device as thermal conduction. From Eq. 2.5.4 the Peltier and Ohmic terms are eliminated and an expression for the temperature drop across the generator can be written

$$\Delta T = T_H - T_C \approx \frac{h(T_H - T_F)}{\left(h + \frac{\kappa}{L}\right)}. \tag{2.8.1}$$

With Eq. 2.5.1, and assuming for simplicity that R_L and R_{int} scale both scale inversely with length, the expression for Q_{gen} in Eq. 2.5.3 can be maximized with respect to length, obtaining:

$$L_{opt} \approx \frac{\kappa}{h}, \tag{2.8.2}$$

$$T_{C,opt} \approx \frac{T_H + T_F}{2}. \tag{2.8.3}$$

While these are approximate expressions, they make it clear that to maximize generated power density about half of the temperature difference between the hot and cold reservoirs driving the generator should be dropped across the heat sink. There is a “thermal impedance match” condition between the generator and the heat sink. This may seem counterintuitive since one normally tries to minimize T_C-T_F by making the heat sink as good as possible. But the point

here is that after choosing the best possible heat transfer coefficient h one can still scale the leg length according to Eq. 2.8.2 to maximize generated power. The optimal length of the couple then depends not only on TE material parameters but also on the quality of the heat sink. A similar result to Eq. 2.8.3 has been presented elsewhere in the limit of small applied temperatures for waste heat harvesting[9].

A more exact computation of the optimum length is possible using the true optimized load for each length. In Figure 19 the power output of the generator is plotted as a function of the device length to find the length maximizing power.

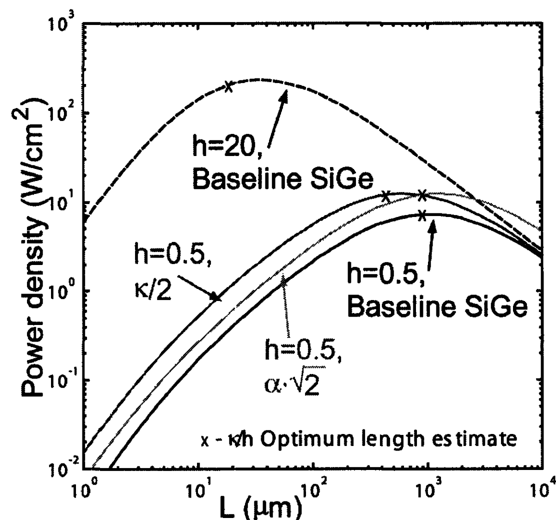


Figure 19. Power density versus leg length for different heat sinks and improvements in selected material parameters.

For baseline SiGe with forced water cooling, the optimal length is approximately 1 mm. Performance at smaller lengths is heat sink limited, and degrades rapidly. To take advantage of the high power densities available at short length scales ($L \sim 30$ μm) a heat sink parameter of around h=20 W/cm^2/K is needed. This performance may be possible using two-phase flow or microchannel heat sinks[48], but has not yet been demonstrated for a TE application. Also shown in Figure 19 are the effects of a factor of 2 decrease in κ and a factor of $\sqrt{2}$ increase in α, each doubling the material's Z. At long device lengths the $\kappa/2$ curve approaches the baseline curve, as more of the applied temperature difference is dropped across the device compared to the

heat sink. In this limit the power output is similar to the expression of Eq. 2.2.4 and is independent of κ.

2.9 Optimum material

To understand how to improve new thin-film TE materials in heat sink limited generators the effects of material improvements on the performance of 100 μm thick baseline SiGe (h=0.5) were investigated. In Figure 20 the effects of equal improvements in the three TE material parameters are shown for both power density and efficiency, where equal refers to the effect on the material's Z. A factor of two increase in σ, a factor of two decrease in κ, and a $\sqrt{2}$ increase in α each increase Z by a factor of 2, and correspond to $Z/Z_{SiGe} = 2$.

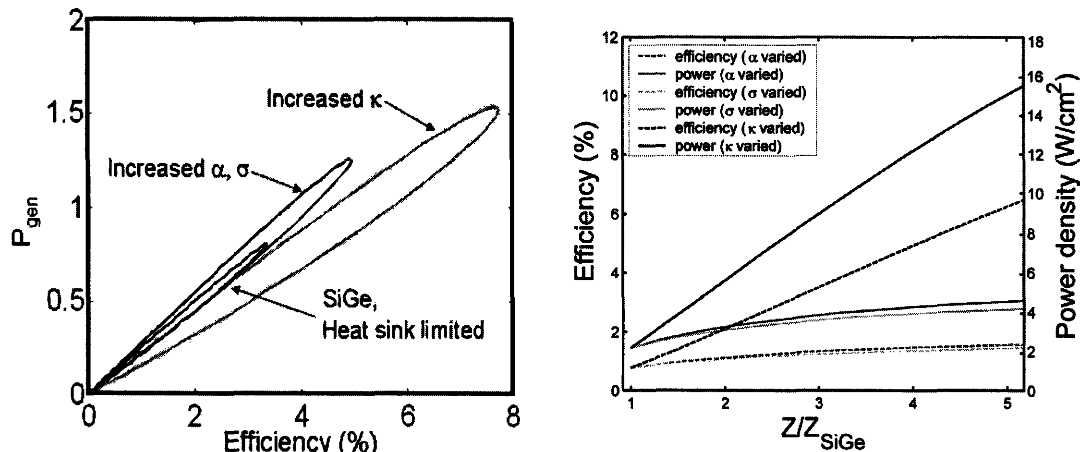


Figure 20. Effects of material improvements on generator performance. The first figure shows the change in power vs. efficiency loops due to increasing in Z by a factor of two for using each of the three material parameters. The second figure extends the analysis to larger changes in Z.

The results of the analysis are quite different from those of the ideal model. Recall that in Eq. 2.2.6 for the ideal efficiency, the material parameters appeared in the lumped parameter Z. The imperfect heat sink and the electrical contact resistance break this simple dependence on Z, and highlight the importance of improvements in κ relative to σ or α. Similarly, in the ideal expression for power density (Eq. 2.2.4) only σ and α appear, but in Figure 20 κ is most important. This result can be understood intuitively. When κ is decreased, not only is the Z increased, but the amount of thermal conduction through the device is decreased, easing the

load on the heat sink. Increasing either α or σ also increases Z, but typically results in more power through the heat sink, partially offsetting the benefit of increased Z.

In general, changing things like the doping[49], polycrystalline grain size[50], or superlattice period[23] can trade off some of the three key TE material parameters (α, σ, κ) against the others. This control is typically used to maximize the lumped TE figure of merit (Z) but we have seen that maximizing Z does not necessarily maximize the power or the efficiency of the generator. The relevant material parameters must be considered individually, and κ is the most important when the heat-sink is limiting performance. For example, in SiGe σ and κ can sometimes be traded off for one another since an increase in doping can increase the electrical conductivity as well as the electronic contribution to the thermal conductivity[51]. Figure 20 makes it clear that to improve efficiency or power in the heat sink limited regime one would choose a somewhat lower doping than that optimizing the Z of the material.

2.10 Conclusions

High-Z thermoelectric thin-films offer the prospect of greatly enhanced thermoelectric power generation if the heat rejection at the cold side of the generator can be managed. Limitations on the maximum heat transfer coefficient of presently available heat sinks and on how thick the TE leg can be grown combine to force present thin-film generators into a heat sink limited regime, characterized by a large temperature drop across the heat sink compared with the device.

In this regime some of the conventional wisdom of thermoelectric generator optimization must be revised. First, the optimum operating current of the generator becomes smaller requiring an increase in the load resistance. The normal impedance matching condition guaranteeing maximum power no longer holds in this regime. Second, the relative impact of the three principal TE material parameters (α, σ, κ) on generator performance is greatly changed. Maximizing the thermoelectric figure of merit Z no longer guarantees optimal efficiency. In the heat sink limited regime changes in the thermal conductivity are the most important for the generated power density; this differs dramatically from the ideal case where the generated power density is insensitive to changes in thermal conductivity.

There appear to be three possible approaches for increasing the generated power density of thermoelectric thin-film generators. First, the heat sink could be improved. New technologies

such as microchannel coolers may be capable of improving heat transfer from the device enough to achieve optimal performance, although other complications are introduced (e.g. the power used to pump the cooler). The focus of Chapter 3 will be on this approach. Second, the length to which high-Z materials can be grown could be increased. If devices can be fabricated with leg lengths in the hundreds of microns optimal performance is possible with presently available heat sinks. The final option is to reduce the thermal conductivity of the TE material. It is not sufficient to simply increase Z. Of the three options, improving the heat sink or the material's κ is the most appealing, since it allows the enhanced power from the short length of the device to be realized.

Chapter 3: High power density heat management

3.1 Overview

This chapter motivates and develops heat management as a strategy for improving thermoelectric generator performance. Chapter 3.2 argues that such an approach is necessary for TE thin-film materials and potentially useful for bulk TE materials. Chapter 3.3 and 3.4 present two models from the literature for heat spreading. The steady-state model introduced in Chapter 3.3 is directly applicable to TE generators and is used in subsequent modeling in this chapter. The dynamic heat spreading model presented in Chapter 3.4 makes use of a handy analytic technique (thermal quadrupoles) that will be used for modeling some of the measurements presented in Chapter 4 and 5. (Appendix 4 introduces thermal quadrupole modeling by the solving for the dynamic heat-transfer though a cylinder, and uses the result to determine the transient errors in meter bar heat flux measurements.) Chapter 3.5 presents a measurement of the heat sink coefficient of a commercially available chilled water cold plate to motivate the discussion of high-performance cooling. Chapter 3.6 introduces the microchannel cooler as a possible solution to the heat management problem and compares the performance of liquid metal relative to water as a coolant. In Chapter 3.7 a model of a TE generator made from commercially available material is integrated with a model of a heat spreader and microchannel cooler. The hydraulic pumping power and other inefficiencies are included in the model allowing the net power production to be calculated. The resulting generator performance is shown to compare very favorably with the performance of a generator with an optimized passive heat sink and with a commercial generator.

3.2 The importance of heat management

There two reasons why working on heat management makes sense as a way to improve thermoelectric generator performance. The first is applicable to thermoelectric elements in general, and is evident from Eq. 2.2.4 for the generated power density, repeated here for convenience:

$$Q_{gen,P} = \frac{(T_H - T_C)^2 \alpha^2}{4\frac{L}{\sigma}} \quad . \qquad (3.1.1)$$

Barring non-idealities such as imperfect contacts and a finite heat sink, the generated power density $Q_{gen,P}$ at the optimum power point scales like 1/L. This means that more power may be generated if a thinner leg size is used or, if preferred, the area of the device may be scaled down while maintaining a constant total power. Eq. 2.8.2 tells us that the optimum element thickness should be set to approximately κ/h (=material thermal conductivity/heat sink coefficient) if three conditions are met:

1) maximizing power density is the goal
2) the electrical contact resistance is small compared to the element resistance
3) the material technology allows the length of the device to be scaled down.

Since optimum power density scales inversely with device length it will increase proportionately with an improved heat sink parameter h.

Many commercial elements could be profitably scaled down to thinner lengths assuming that a practical way to deal with the corresponding increase in heat power density could be found. The main reason why commercial devices are not thinner is the heat-sink limitation discussed in Chapter 2. A typical package thermal conductivity for a commercial thermoelectric generator is 0.024 W/cm/K (HZ-2 from Hi-Z Technology, shown in Figure 21). If one assumes that the user of such an element mounts the generator in direct contact with a solid metallic hot side and uses a heat sink whose heat transfer coefficient is consistent with passive water cooling ($h \sim 0.05$ $W/cm^2/K$), one finds an optimal length of 0.52 cm, according to Eq. 2.8.2. This is close to the actual Hi-Z generator thickness of 0.508 cm.

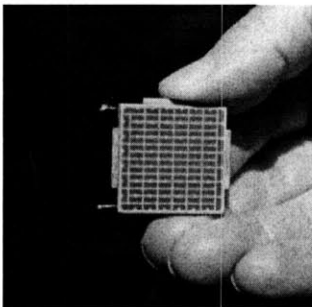


Figure 21. HZ-2 thermoelectric module from Hi-Z Technology, Inc.

Barring limitations on the heat sink, a limit to how much the element leg length can be profitably reduced is set by the electrical parasitic contact resistance. When the contact resistance of the element R_C becomes larger than the bulk resistance, decreasing the device length no longer increases power density (see Eq. 2.5.3 with Eq. 2.5.1). Decreasing the length past this point never makes sense, since efficiency also drops with decreasing L (Eq. 2.5.2). Even if the

contact resistance could be made negligible (a dubious prospect for the hot-pressed Bi_2Te_3 material used for the HZ-2 or for the refractory metal contacts needed for high temperature contacts) the mechanical stability of a thin multi-crystalline or a powder metallurgical element might also be a limiting factor at small lengths. In Chapter 4 the performance of a commercial powder metallurgy material is measured as its length is scaled down. If one could supply sufficient heat sinking, these devices could be scaled down to below 100 μm before contact resistance seriously limited performance.

The second reason why heat management is important for TE power generation is that that the thin-film materials presently having the best measured thermoelectric material properties can only be grown using molecular beam epitaxy (MBE) or other epitaxial growth methods, severely limiting the maximum practical thickness. As argued in Table 1 of Chapter 2, elements made from these materials are presently heat sink limited. Improving the heat sink will therefore dramatically increase both efficiency and power density.

These materials typically have structures on the nanometer scale that are specifically tailored to improve thermoelectric performance, so that growth must proceed at a very slow rate. For MBE-grown thin-film elements the maximum thickness is typically limited by physical constraints or cost. An example of the former is the pseudomorphic (strained) growth of quantum well or superlattice structures. The buildup of strain may limit the growth to a critical thickness depending on the degree of lattice mismatch and the details of the growth kinetics[52]. This is less of a problem for strain-relaxed or lattice-matched growths although practical limitations on source stability can still set a maximum attainable thickness. However, extrapolating the epitaxial growth costs for TE elements from those associated with microwave transistors or optical components is probably not valid. Unlike a typical high-electron mobility transistor or quantum-well laser, high power TE elements are inherently large-area (since power scales with area). The growth thickness required for a practical TE device (due to the heat-sink limitations discussed in Chapter 2) is also greater, increasing material costs and requiring more reactor time. However, it is also likely that the stringent requirements for low defect density needed for high-mobility transistors and optoelectronics can be relaxed somewhat for thermoelectric devices, perhaps allowing for faster growth and increased yield. It is fair to say that right now thick MBE growth of thermoelectric materials is prohibitively expensive. As a result, at the present time we are aware of only one example of an epitaxially grown TE element longer than

20 μm in length[21]. The epitaxial materials studied in this thesis were all grown using MBE, and most were under 5 μm in length.

Working on improving the heat transfer out of the element makes sense for both advanced nanostructured thin-films and presently available commercial materials. In the former case, improved heat transfer could enable the (inconveniently) thin elements to be usefully employed. In the latter case, it enables existing bulk materials to be scaled down to increase power, broadening the design space for thermoelectric power generation.

3.3 Heat spreading for thermoelectric power generation

The technique of heat spreading is commonly used in the microelectronics industry to manage the large thermal gradients in high-power optoelectronic devices and circuits. In light of the difficulty of extracting heat from the cold side of thin-film thermoelectric generators discussed in the previous section and in Chapter 2, it is worthwhile to examine heat spreading as a possible means of performance enhancement.

The idea behind heat spreading for thermoelectric power generation is to vary the area of the TE element relative to the heat sink in order to match the total heat though the elements to the heat that can be realistically managed by the heat sink. Heat spreading can therefore be understood as a means of thermal impedance matching. From Eq. 2.8.2, the optimal power generation occurs approximately when the thermal resistance of the TE element is equal to that of the heat sink. Here we will model the heat sink as a slab of some substrate serving as a heat spreader in series with a plane of known heat transfer coefficient h, with units of [W/cm^2/K].

If the thin-film is allowed to have a smaller area than the substrate, the thermal resistance of the film is still approximately

$$R_{th,film} = \frac{L_{film}}{\kappa_{film} A_{film}}. \qquad (3.2.1)$$

However, with heat spreading from the film into a substrate the thermal impedance of the combined substrate and heat sink becomes more complicated because the heat transfer is no longer one dimensional. The situation is analogous to the problem of determining the electrical resistance of a 3D solid of arbitrary shape. In fact the calculations are formally equivalent; the calculation of the resistance in both cases involves the solution of the Poisson equation with von

Neumann or Dirichlet boundary conditions, with temperature playing the role of voltage and heat flux playing the part of electrical current. There is one caveat: one typically has to be more careful with the calculation and measurement of thermal resistances than with their electrical analogues because most practical cases of interest in the electrical domain involve metal contacts that define the geometric boundaries of the object at nearly perfect equipotentials, while for the temperature case such ideal boundaries (isotherms or isopowers) are not guaranteed. The definition of lumped thermal impedance for a 3D object only makes rigorous sense between well-defined 2D boundaries of constant temperature (or heat flux). For modeling purposes, this requirement for defining the thermal impedance is often relaxed. Thermal impedance is then defined between geometrically convenient reference planes whose total heat flux and average temperature are defined.

The boundary conditions for the heat spreading at the cold side of a thermoelectric generator can be well-approximated by a constant heat flux boundary condition (from the thermoelectric material) into the substrate and a fixed heat transfer coefficient h to model a heat sink at the bottom of the substrate. Convective heat transfer from the sides and top of the substrate can be neglected. These assumptions are sufficient to calculate the thermal resistance using an analytic model given by Lee et al. [53] for the substrate thermal spreading resistance:

$$R_{Th} = \frac{L}{A\kappa} + \frac{1}{2a\kappa\sqrt{\pi}}(1-\varepsilon)^{3/2}\Phi_C, \tag{3.2.2}$$

$$\Phi_C = \frac{\tanh(\lambda_C\tau) + \frac{\lambda_C}{B}}{1 + \frac{\lambda_C}{B}\tanh(\lambda_C\tau)}. \tag{3.2.3}$$

The input variables here are:

$$A = \pi b^2 \qquad B = \frac{hA}{\pi\kappa b} \qquad \varepsilon = \frac{a}{b} \qquad \tau = \frac{L}{b} \qquad \lambda_C = \pi + \frac{1}{\varepsilon\sqrt{\pi}} \qquad B = \frac{hA}{\pi\kappa b}$$

The model assumes a cylindrical geometry to perform an expansion in orthogonal functions. For this geometry "a" is the radius of the uniform heat source and "b" is the radius of the substrate (Figure 22).

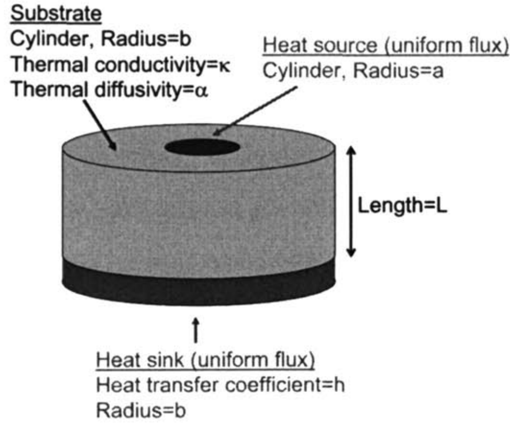


Figure 22. Geometry and parameters used for heat spreading calculations. A concentric cylindrical geometry is assumed to allow for a closed form solution in terms of orthogonal functions.

Lee et al. also tested that the model works well for other geometries of interest (e.g. square-shaped TE elements), so we adopt this analytical model for further studies on device optimization.

A test of how the heat spreading model (which assumes a circular heat source and substrate) compares to a numerical solution of a rectangular heat source and substrate was performed. The thermal resistance was calculated from the analytic formulas of 3.2.2 and 3.2.3 and separately from the results of a 3-D finite element simulation (FEMLAB). The FEMLAB spreader consisted of a square $A_{bot}=\pi b^2=1$ cm^2 area of 500 μm thick silicon wafer. On the bottom of the slab a boundary condition consistent with forced water cooling (h=0.5 W/cm^2/K) was imposed. Centered on the top of the slab was a square-shaped heat source of 0.2 cm on a side ($A_{top}=\pi a^2=0.04$ cm^2) to model the footprint of a thin-film TE mesa. An identical area heat source was used for the analytic calculation. The input heat power density in both cases was assumed to be $Q_{top}=100$ W/cm^2. A thermal resistance may be defined in this case as:

$$R_{th} = \frac{\langle T^{top} \rangle - \langle T^{bot} \rangle}{P_{in}} = \frac{\frac{1}{A_{top}} \iint_{A_{top}} T \cdot dA_{top} - \frac{1}{A_{bot}} \iint_{A_{bot}} T \cdot dA_{bot}}{Q_{top} A_{top}} . \qquad (3.2.4)$$

Here $\langle T^{top} \rangle$ and $\langle T^{bot} \rangle$ are the area-averaged mean temperatures on the top (heat source of area a) and on the bottom (heat sink of area b) of the spreader, and Q_{top} is the heat passing through the spreader. Using the finite-element model the temperature is solved for everywhere in the bulk and then the appropriate numerical integrals are taken over the input and output

planes according to Eq. 3.2.4. One finds R_{th}=2.20 K/W from the analytic formula and 2.25 K/W from the FEMLAB model (the results of which are shown in Figure 23), corresponding to better than 3% agreement in spite of the geometry difference.

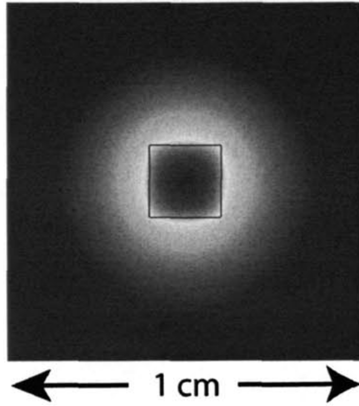


Figure 23. Finite element solution for the temperature distribution at the top of a heat spreader using FEMLAB. The color scale is proportional to temperature above room temperature. The isotherms are nearly circular, even though the heat source is square-shaped.

The temperature isotherms around the square-shaped heat source in Figure 23 appear quite circular, supporting the conclusion of Song et al. [53] that an arbitrarily shaped source generates numerically similar thermal resistances to a circular source of the same area and broadening the applicability of their analytic model.

We may easily incorporate heat spreading into the design of optimized TE generators by adding the thermal spreading resistance into the model in series with the heat sink parameter which is already present. In what follows an example is worked through. The TE element is assumed to be a thin-film (4 μm thick), and the material parameters for the simulated TE were α=-247 μV/K, κ=0.0393 W/cm/K, and σ=485 1/Ω/cm. These values are for SiGe, but are also similar to those expected for the nanodot ErAs material that will be discussed in Chapter 4. The cold side heat spreader is assumed to be a 200 μm of CVD diamond with κ=10 W/cm/K[54]. To remove the heat from the bottom of the spreader, active water cooling is assumed (heat transfer coefficient h=0.5 $W/cm^2/K$, water temperature 300 K). Because an ideal temperature boundary condition on the hot side of the small area thin-film is unphysical, a heat concentrator is modeled using a 1 mm thick copper wall (κ=3.91 W/cm/K), one side of which is maintained at 923 K using an ideal temperature boundary condition. Other parasitic effects, including thermal and electrical

interface resistances and parasitic heat conduction across the gap between the hot and cold side spreader are ignored. Figure 24 shows the results of the simulation. Both the generated power density and the fraction of the applied temperature dropped across the TE element are plotted. The x-axis of both plots is the ratio of the area of the thin-film TE material to that of the heat concentrator/spreader, a key design parameter. From Figure 24 it is clear that for the model assumed here adding a heat spreader and concentrator is a useful design technique for increasing the generated power density assuming the area ratio is chosen well. However, it is also clear that for any area ratio (including that which optimizes the power density), less than 50% of the applied temperature difference is actually dropped across the TE element. Even at the optimum power area ratio the device is heat-sink limited.

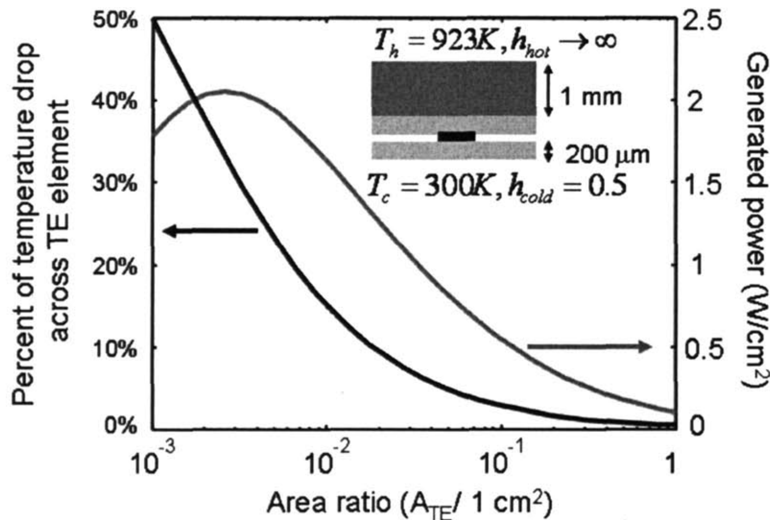


Figure 24. The effect of heat spreading on the temperature difference across a thin-film element and on the generated power.

When additional non-idealities associated with the heat spreading technique are accounted for the temperature drop across the TE element can decrease further. For many practical waste-heat recovery applications a gas must be used as either the hot or cold reservoir, reducing the heat transfer relative to that assumed for Figure 24 by well over an order of magnitude. For thin-film elements hundreds or thousands of series interconnections between the elements are needed to impedance match to a reasonable load for practical power generation. The parasitic electrical resistances introduced by this scheme will become more difficult to manage when the individual elements are connected with long wire contacts, as is presumably necessary for

smaller area ratio designs. Also, in the simulation here the effects of parasitic radiative and convective heat transport across the thin (in this case 4 μm) gap between the hot and cold side of the element have been ignored. This is optimistic, especially for the smaller area ratios in Figure 24 where the area for this parasitic heat transfer is orders of magnitude larger than the element area. All of these parasitic issues reduce the efficacy of the heat spreading approach lessening the percentage of the total applied temperature which is usefully dropped across the element.

For many cases of practical interest, there would appear to be no simple solution to the serious thermal impedance mismatch issue highlighted by Table 1 in Chapter 2.4. Until significant breakthroughs are made it appears that generators attempting to use some of these thin-film materials will be forced to operate to some degree in the heat-sink limited regime introduced in Section 2.4.

3.4 Heat spreading in the transient regime

The heat spreading model of Lee et al. described in Section 3.3 is meant to describe the steady-state thermal impedance of a heat spreader. This is the relevant calculation for most power generation applications but it is also useful to have a model for heat spreading that is applicable to non-steady-state situations. The method of thermal quadrupoles is a general technique for finding dynamic thermal impedances in complex geometries[55]. The method works by subdividing the heat transfer domain of interest into geometrically simple regions in which the time dependent heat equation can be written using coordinates that take advantage of any available symmetry. One handles the time dependence by transforming each resulting differential equation into the complex frequency domain using the Laplace transform, and then solving the resulting partial differential equation using standard methods (e.g. orthogonal function expansion in the coordinate system of choice) subject to fixed (Laplace transformed) temperature and flux boundary conditions on the two surfaces which define the thermal impedance of each domain. The result can be expressed as a 2x2 matrix for each region relating the temperature and flux at one boundary and the temperature and flux at the other. The thermal impedances of the subdivided domains, now described by matrices, can then be cascaded together using matrix multiplication to rebuild the dynamic thermal impedance of the original complex structure. An example is useful to clarify the method. In Appendix 4, the thermal quadrupole of a uniform cylindrical conductor is found and it is used to calculate an

important potential source of error of the meter bar heat flux measurements that are presented in Chapter 4.

We return to the cylindrical geometry heat spreader shown in Figure 22. The thermal quadrupole matrix for the geometry relates Laplace transforms of the average temperature $T_{top}(s)$ and total heat flux $P_{top}(s)$ (in Watts) at the heat source to the Laplace transforms of the average temperature $T_{bot}(s)$ and the total heat flux $P_{bot}(s)$ at the bottom of the substrate:

$$\begin{bmatrix} T_{bot}(s) \\ P_{bot}(s) \end{bmatrix} = M_{spread}(s) \begin{bmatrix} T_{top}(s) \\ P_{top}(s) \end{bmatrix}.$$

The boundary conditions are a constant heat flux at the heat source and a heat transfer coefficient h at the back of the spreader. The general method for finding the matrix M_{spread} has been shown elsewhere[55]. The matrix for this geometry and boundary conditions is given by

$$M_{spread} = \begin{bmatrix} \cosh\left(L\sqrt{s/a_{diff}}\right) & -\dfrac{\sinh\left(L\sqrt{s/a_{diff}}\right)}{\kappa A\sqrt{s/a_{diff}}} - \cosh\left(L\sqrt{s/a_{diff}}\right)\Sigma \\ -\kappa A\sqrt{s/a_{diff}}\sinh\left(L\sqrt{s/a_{diff}}\right) & \cosh\left(L\sqrt{s/a_{diff}}\right) + \kappa A\sqrt{s/a_{diff}}\sinh\left(L\sqrt{s/a_{diff}}\right)\Sigma \end{bmatrix},$$

where $\Sigma = \sum_{n=1}^{\infty} \dfrac{4J_1(\xi_n a)}{\kappa A a^2 \gamma_n J_0(\xi_n b)} \dfrac{\frac{\kappa\xi_n}{h} + \tanh(\gamma_n L)}{1 + \frac{\kappa\xi_n}{h}\tanh(\gamma_n L)}$.

As in the previous section, a is the radius of the heat source, b is the radius of the substrate, L is the length of the substrate, $A=\pi b^2$ is the area of the substrate, κ is the substrate thermal conductivity, and a_{diff} is the substrate diffusivity. J_1 is the Bessel function of the first kind of order one, and J_0 is the Bessel function of the first kind of order zero. The coefficients ξ_n and γ_n arise from matching the orthogonal Bessel functions of the solution to the specified boundary conditions. They are defined by the following two eigenvalue equations:

$$J_1(\xi_n b) = 0$$

$$\gamma_n^2 = \xi_n^2 + s/a_{diff}$$

Efficient numerical routines exist for calculating the zeros of the Bessel function, and we have found that the first few hundred are sufficient for accurate computations.

It is shown in Appendix 4 how the quadrupole matrix describing a passive system can be used to find an equivalent thermal circuit model for the heat spreading problem (shown in Figure 25).

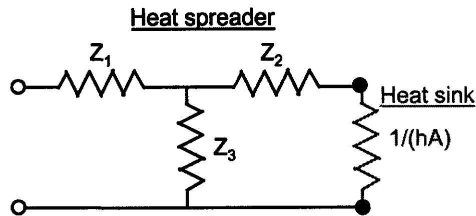


Figure 25. The equivalent thermal impedance network used to model the heat spreader. A thermal resistance modeling the heat sink on the bottom of the heat spreader has also been included.

Any passive, lossless, linear, reciprocal two-port can be described by a three resistor impedance T-network (or a Π network, which we will not need to use), like the one shown in Figure 25. 'Passive' and 'lossless' in this context mean that heat is conserved in the device, and must all either enter or exit through the two defined terminals. 'Linear' refers to the linearity of the Fourier law governing heat conduction, that is, the linear response of heat flow to an applied temperature gradient. 'Reciprocal' in this context means the following: If some temperature T is applied to the input port and the output port is thermally shorted out, some heat current P will flow through the output port. If we now start over and connect the temperature T to the output port and thermally short the input port, reciprocity requires that the heat flowing though the input port will again be equal to P. Interestingly, neither physical symmetry nor macroscopic time-reversal symmetry is required to ensure reciprocity. To ensure reciprocity, it is sufficient that the thermal conductivity tensor relating the heat flux to the temperature gradient (in the three-dimensional Fourier law) be real and symmetric. It is not necessary that it be isotropic. (It is interesting that the symmetry of the thermal conductivity tensor can apparently be derived in a non-trivial manner from the Onsager reciprocity relations[56], which are in turn a consequence of microscopic irreversibility.) The construction of the T-network from the two-port matrix is discussed in more detail in Appendix 4.

In the network we include a terminating resistor modeling the heat sink at the bottom of the heat spreader. (We are not double counting; the heat sink coefficient as defined in Chapter 2.4 shows up in the spreader calculation only as a boundary condition affecting the heat transport in the spreader, not as a thermal resistance per se.) The impedances Z_1, Z_2, and Z_3 model the heat spreader and substrate conduction and they can be calculated from the quadrupole matrix M_{spread}. Using the formulae presented in Appendix 4 one finds:

$$Z_1 = \Sigma + \frac{1}{\kappa A\sqrt{s/a_{diff}}\tanh\left(L\sqrt{s/a_{diff}}\right)} - \frac{1}{\kappa A\sqrt{s/a_{diff}}\sinh\left(L\sqrt{s/a_{diff}}\right)}$$

$$Z_2 = \frac{1}{\kappa A\sqrt{s/a_{diff}}\tanh\left(L\sqrt{s/a_{diff}}\right)} - \frac{1}{\kappa A\sqrt{s/a_{diff}}\sinh\left(L\sqrt{s/a_{diff}}\right)}$$

$$Z_3 = \frac{1}{\kappa A\sqrt{s/a_{diff}}\sinh\left(L\sqrt{s/a_{diff}}\right)}$$

From the circuit model of Figure 25 it is straightforward to calculate the thermal impedance seen at the top of the heat spreader:

$$Z_{th} = Z_1 + Z_3 \,||\, \left(Z_2 + \frac{1}{hA}\right).$$

Using this expression the thermal resistance of the heat spreader may be calculated as a function of frequency. To test this expression, we modeled heat spreading from a source of heat of radius a=100 μm into a silicon substrate (b=1000 μm, L-500 μm, α=0.8, κ=1.3) backed by a heat sink (h=0.6 W/cm^2/K). The results are shown in Figure 26.

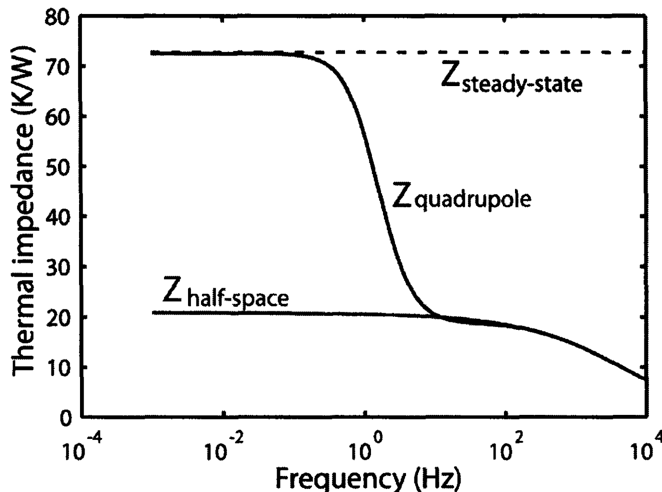


Figure 26. A comparison of various analytic models for the cylindrical thermal spreading resistance. The dotted 'steady-state' line models the steady state thermal resistance model of Lee, Song, Van Au, and Moran discussed in the previous section. The 'half-space' line is the solution for the dynamic spreading resistance into an infinite half-space. The 'quadrupole' line is the solution of the dynamic quadrupole model presented in this section.

We compared the thermal impedances calculated using three models. The first model (the dotted line marked '$Z_{steady-state}$') is the steady-state spreading resistance of Lee et al. presented in

the previous section. It is valid only for zero frequency. The line marked '$Z_{half\text{-}space}$' models the heat spreading into an infinite half space (a simpler idealization of the heat spreading problem than the one worked out in this section). The line marked '$Z_{quadrupole}$' is the impedance calculated using the quadrupole formulae developed in this section. The quadrupole result agrees with the steady-state thermal resistance at low frequencies up until a certain break-point frequency. We can estimate the break-point by calculating $f_{3db} = \frac{1}{2\pi R_{th} C_{th}}$, where R_{th} is the steady-state thermal impedance and C_{th} is the heat capacitance of the entire substrate (since in our problem, at low frequencies the heat sink resistance dominates and the entire substrate is heated up):

$$R_{th} \approx 72\frac{K}{W}, \quad C_{th} = \frac{\kappa}{a_{diff}} LA \approx 0.0026\frac{J}{K}$$

$$\rightarrow f_{3db} \approx 0.85\,\text{Hz}$$

It is also evident from Figure 26 that the quadrupole calculation approaches the half-space result at high enough frequencies. This can be understood in terms of the frequency dependent thermal penetration depth, which is equal to $d_{thermal} = \sqrt{\frac{a_{diff}}{4\pi f}}$ [57]. In silicon, one calculates that once the thermal excitation frequency is greater than about 25 Hz the thermal wave can no longer "see" the back of the wafer 500 μm away from the heat source, and so the full system becomes well approximated by the idealized infinite half-plane solution.

While thermoelectric power generation is typically a steady-state process, the full frequency dependent quadrupole solution will be of use to us later when we discuss some of the transient measurements used to characterize thermoelectric materials.

3.5 Experimental characterization of a heat sink

It is important to develop a system to characterize the heat transfer coefficient of a heat sink experimentally. By doing so, the usefulness and accuracy of the heat transfer coefficient as a descriptive tool for heat sink applications may be checked. Several techniques are available[58]. The one we chose is to use a bar of material of known thermal conductivity (a "measure bar" or "meter bar") to inject heat into the heat sink while monitoring the temperature at regularly spaced points along the bar. This is depicted schematically in Figure 27 (the three thermocouples measure temperatures T_1, T_2, and T_3.) The thermocouples are placed in holes which extend to the center of the bar. The bar is made of 6061 aluminum with a thermal

conductivity of 1.8 W/cm/K, and has a square profile one inch on a side. Heat is supplied using a PID stabilized resistive heater sunk in an aluminum block and pressed against the top of the bar.

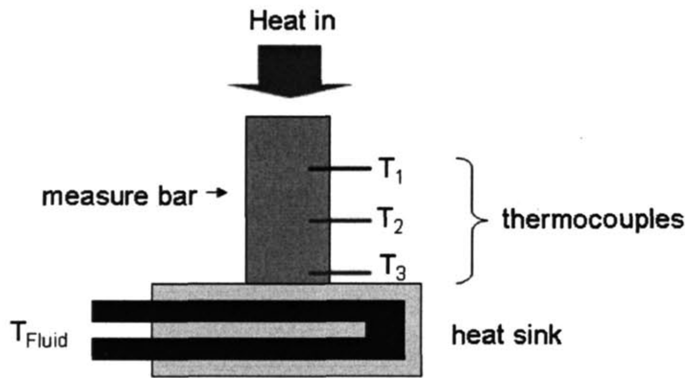


Figure 27. Basic geometry for a measure bar used to measure the heat transfer coefficient into a heat sink.

The setup was used to characterize the heat transfer coefficient of a commercial chilled water heat sink (LC Cold Plate, Solid State Systems) using a measurement similar to that used in the commercial literature [59]. The temperature T_1 was used to stabilize the heater input temperature and the heat transfer coefficient h was determined from the temperature difference between T_2 and T_3, the measure bar area (A), the thermal conductivity of the aluminum (K), the distance between the T_2 and T_3 thermocouples (L), and the measured temperature of the coolant water (T_{Fluid}) according to the formula:

$$\frac{K(T_2 - T_3)A}{L} = h(T_3 - T_{Fluid})A \quad (3.4.1)$$

The results are shown in Figure 28. As a check that the measurement was working, the hot side of the meter bar was stepped up to three different heater temperatures. After an initial thermal transient the heat transfer coefficient was measured to be 0.66 W/cm^2/K, consistent with the range expected for forced water convection (e.g. Table 1).

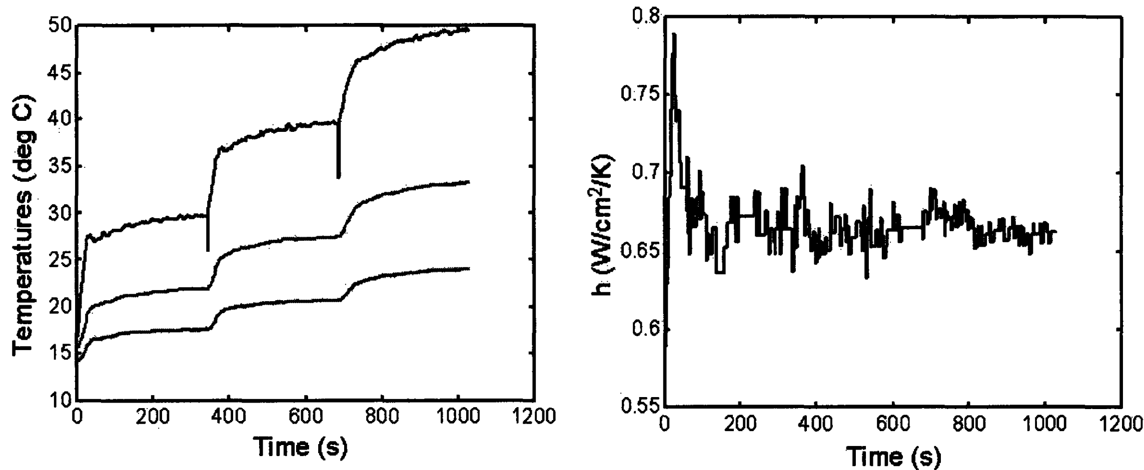


Figure 28. Experimental determination of the heat transfer coefficient for a chilled water cold plate. Raw thermocouple data is shown at left, and the inferred heat transfer coefficient is shown at right.

There are a few points that should be mentioned regarding the interpretation of heat transfer coefficients obtained from this type of experiment. The heat transfer coefficient measured in this manner may actually depend on the area of the measure bar since the area ratio of the measure bar to the actual heat sink will change the thermal spreading resistance according to the results discussed in Section 3.2. This is unavoidable and simply means that the diameter of the measure bar should be selected to match the physical system whose performance is to be characterized. Second, the heat transfer coefficient may generally depend somewhat on the temperature at the sink, although that was not the case in Figure 28. An important example for a water cooled sink is the onset of boiling and the resulting two-phase flow, which completely alters the heat transfer coefficient (as well as the hydraulic performance). Finally, when determining the heat transfer in the meter bar one must wait a certain amount of time (related to the heat capacitance of the meter bar and the thermal resistance of the device) before the measurement is valid. This is evident from the transients visible in the measurement of h. This constraint is investigated in Appendix 4 using the analytic methods developed in the preceding section.

Other potential sources of error in this measurement are the assumed value of the thermal conductivity of the 6061 aluminum (from a table of published values at room temperature) and from the assumption of uniform one-dimensional conduction through the bar. Any convective or radiative heat transport from the side of the meter bar will preferentially remove heat from the top of the bar, skewing the linear temperature profile and introducing some error into the results.

For the measurement described above, 3D-finite element modeling of the heat transfer from the bar revealed that the effects were negligible. But in a measurement using a thinner bar (necessary to look at the effective heat transfer coefficient of a smaller heat sink or heat transfer through small samples) or at higher temperatures the error can become more significant. For this reason doing the experiment in vacuum and with the smallest feasible temperature difference is helpful to ensure good accuracy. Leaving sufficient distance between the constriction and spreading that occurs as heat enters or leaves the bar and the thermocouples that monitor heat flux is a good idea. It is also good practice to use more than two thermocouples in the meter bar so that the nonlinearity in the temperature profile that indicates a parasitic heat loss can be detected. Subsequent meter bars used in this thesis make use of these improvements.

3.6 Microchannel heat sink – channel simulations

The previous sections have highlighted the importance of the heat sink in the design of a high-power density thermoelectric generator and laid the groundwork for modeling and measuring heat transfer into a heat sink. Here we discuss microchannel cooling as a means of increasing the heat transfer coefficient (referred to by h) directly. We begin by developing a model of the fluid and heat transport in a microchannel and then we apply the results to a potentially useful idea that has been suggested in the literature, the use of liquid metals as coolants [60].

Heat management of semiconductor devices in general is a topic of increasing engineering importance on both a chip and device level, for applications much broader than thermoelectric power generation [61]. For conventional silicon technology the continuing increase in the speed and area density of transistors predicted by Moore's law (and the related increase in metal interconnect layers) has made power dissipation a more and more critical consideration in modern circuit designs. On the device level, many power transistors and high power semiconductor lasers can generate large spot heat fluxes (> 1000 W/cm^2) which must be directed away from the device for good performance. By way of comparison, heat fluxes in the hundreds of Watts/cm^2 range are expected in thin film thermoelectric devices.

Of the technologies considered in Table 1 in Chapter 2, the only one capable of delivering the high cooling power densities needed for thin-film thermoelectrics is the microchannel heat sink. The microchannel heat sink concept was originally pioneered by Tuckerman and Pease as a

solution to VLSI heat problems in 1981[62]. They fabricated a 1 cm^2 array of 50 μm x 302 μm channels by anisotropically etching a silicon substrate. Using a pressure of 31 psi to force water through the heat sink they achieved a very low thermal resistance of 0.09 K/W/cm^2. Since then many subsequent efforts have yielded similarly enhanced cooling performance in other material systems and geometries[63]. A schematic of a "canonical" microchannel heat sink is shown in Figure 29 [64].

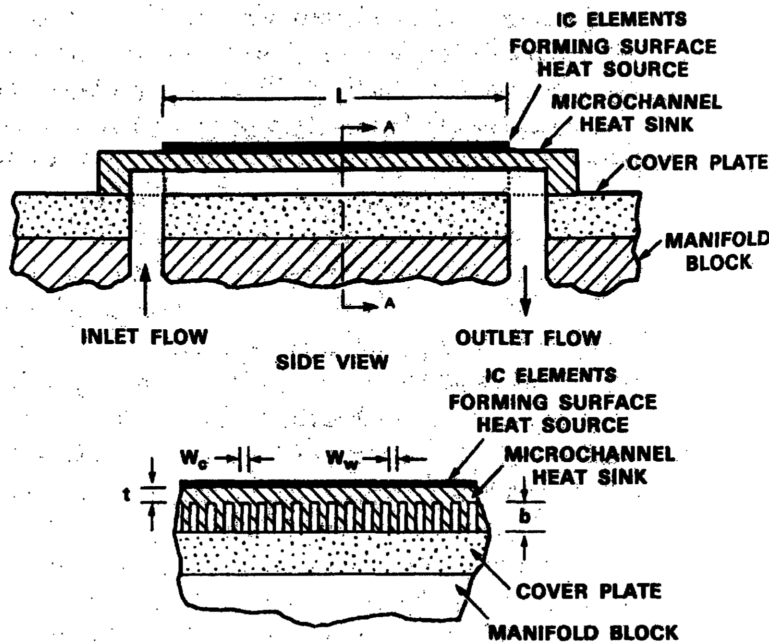


Figure 29. Schematic diagram of microchannel cooler geometry (from Phillips, 1990).

Microchannels are typically cut or etched into a plate made of a material with a high thermal conductivity (copper, silicon, or even diamond) onto which the heat-producing device is grown or bonded to minimize thermal interface resistance. The channels are then capped with another plate (whose thermal performance is less crucial) to form the closed channels. A manifold and a pump are designed to move the working fluid (and heat) away from the chip and into some other heat exchanger or thermal reservoir. There are two key parameters of interest for thermoelectric generation. The first is the system heat transfer coefficient h_{sys} (W/cm^2/K) determining how much heat flux can be removed per unit area of the device. The second is the required hydraulic power, characterizing how energetically expensive it is to push water through the heat sink. The latter is obviously a key concern since our goal at the end of the day is to generate power.

There are several reasons for the very low thermal resistances attainable with microchannel cooling. First, the channels can be machined into the substrate material and placed in very close proximity to the thermoelectric device. In terms of Fig. 3 above, the distance t between the surface of the cold plate and the tops of the channels can be made very small. Second, by using deep channels the water-semiconductor interface area is made very large compared to the area of the device. Thus the heat transfer coefficient into the water does not generally limit h_{sys}; there is an additional enhancement in h_{sys} due to the increased area afforded by the corrugated surface of the solid-liquid interface. Finally, for a given fluid flow velocity the heat local transfer coefficient h [W/cm^2/K] characterizing the heat removal at the solid-liquid interface becomes larger for microchannels (at a price of higher pump head pressures).

One may understand the main reason why microchannels enhance the channel wall heat transfer coefficient by considering laminar fluid flow between two closely spaced parallel plates, the simplest geometry for channel flow. Under laminar flow conditions the cross-sectional velocity profile will be parabolic in shape; maximum in the center of the channel and zero on the channel walls (assuming a no-slip boundary condition at the walls). If the temperature of the plates is raised above that of the water, heat will flow from the plates into the water in the parallel plate channel. The temperature profile in the channel (for fully developed flow and thermal conditions) will resemble the opposite of the velocity profile; in the center of the channel where the flow velocity is highest and heat is rapidly carried down the channel by the flowing water, the water temperature will be lowest. Suppose we decrease the plate separation while keeping the mean fluid flow velocity fixed. As the plate separation is narrowed while the flow velocity is held constant, the parabolic profile is forced to become more "peaked" to maintain the same mean flow velocity and the gradient in the velocity becomes steeper near the channel wall. As a result, the gradient in the thermal profile near the channel wall becomes steeper as well. Since the heat transfer into the channel is exactly to the heat flux at the channel walls, the channel wall heat transfer coefficient has increased. This enhancement in the heat transfer coefficient into narrower channels is also sometimes described as a thinning of the "thermal boundary layer" in the channel. The concept of a boundary layer is even more appropriate for turbulent flow. The parallel plate discussion has so far focused on laminar flow conditions, but as some point, the channel dimensions become thin enough (and the shear forces on the fluid large enough) that the flow breaks up and becomes turbulent. In this case, the flow velocity profile flattens out except for a thin boundary layer with a very steep (time averaged) velocity

gradient that maintains the no-slip zero flow velocity at the wall. The mathematical description of the behavior of the turbulent boundary layer is rather more involved than the laminar case, but at higher flow velocities, it also becomes thinner. The turbulence also has the desirable (from the standpoint of cooling) effect of mixing the flow in the channel, helping to increase heat transfer. This comes at a cost of more internal heating through frictional losses in the fluid and higher power required to pump the flow.

A simulation generating the fully-developed laminar temperature and flow distributions in a rectangular channel is described here. This is also a canonical analytical problem in fluid mechanics, allowing the validity of the finite-difference simulation to be checked against the exact solution. The relative performance of two fluids, water and liquid metal, is then investigated in conjunction with the microchannels.

For this simulation we will assume fully developed laminar fluid flow. "Fully developed" means that the part of the channel to be simulated is far from the entrance to the channel, so that the flow's velocity and temperature distribution is no longer changing. Laminar flow assumes that the individual fluid elements in the channel follow smooth streamlines, allowing us to use Newton's law of shear stress to directly calculate the flow. We are motivated to consider this flow regime because for typical channel dimensions the turbulent regime requires inconveniently high pump pressures, an important consideration if the microchannels are used in conjunction with thermoelectric generation.

The condition for laminar flow in a pipe is that dimensionless Reynolds number (Re) averaged across the channel area is less than 2300 [65]. (Experimental results suggesting a faster onset of turbulence in microchannels are common but they have proven somewhat irreproducible due to their sensitivity to boundary effects and channel wall roughness.) In the context of pipe flow, the Reynolds number is given by $Re = \frac{\rho u D_h}{\mu}$. Here ρ is the fluid's density [kg/cm^3], u is the mean fluid flow velocity in the channel [cm/s], and μ is the fluid's dynamic viscosity [kg/cm/s]. The hydraulic diameter D_h [cm] is given by 4 times the area of the pipe divided by the perimeter of the pipe. Expressed in terms of D_h, analytic results for fluid flow through a specific pipe cross-section can sometimes be used to model a different pipe cross-section. For a circular pipe, the hydraulic diameter is simply the pipe diameter in the usual sense. There are a few useful dimensionless parameters that are helpful in discussing microchannel cooling: the Prandtl

number $\left(Pr = \frac{\mu c_P}{k}\right)$ that is the ratio of momentum diffusivity to heat diffusivity, the Nusselt number $\left(Nu = \frac{hD_h}{k}\right)$ describing the relative importance of convective to conductive heat transfer in lateral heat flow, and the Darcy friction factor $\left(f_D = \frac{2\tau_s}{\rho u^2}\right)$ which can be used to calculate the heating associated with the fluid flow due to viscous effects. Here k is the fluid's thermal conductivity [W/cm/K], h is the heat transfer coefficient into the channel [W/cm^2/K], c_p is the fluid's specific heat [J/K/kg], and τ_s is the local wall shear-stress [N/cm^2]. The analysis we will do assumes known thermophysical properties and a mean channel flow velocity $\bar{u}$ for the fluid, and then calculates the mean heat transfer coefficient at the wall, $\bar{h}$ (or equivalently, $\overline{Nu}$), as well as the hydraulic power needed to drive the flow (related to $\bar{f}_D$).

For turbulent flow or flow in transition between laminar and turbulent flow, accurate analytical modeling is more difficult. One typically resorts to a "correlation," a parametrized model for the dimensionless numbers characterizing the flow and energy transport obtained from parametrically varied experiments or full fluid dynamical simulations. Correlations for Nu are available for both laminar and turbulent flow, for example[66],

$Nu_{turb} = \frac{(f_D/8)Re\,Pr}{1.07 + 12.7\sqrt{f_D/8}\left(Pr^{2/3}-1\right)}$. Here we would calculate Nu directly from the flow characteristics.

Starting with Newton's law of shear stress and the pressure differential along the flow direction an equation for the flow velocity in the channel can be found by summing all of the forces on a differential volume of fluid, resulting in equation similar to the Poisson equation of electrostatics.

$$\frac{dP}{dx} = -\mu\left[\frac{d^2}{dy^2} + \frac{d^2}{dz^2}\right]u(y,z) \tag{3.6.1}$$

Here $\frac{dP}{dx}$ is the pressure drop occurring across the differential slice of the channel in the direction of flow. The directions y and z are perpendicular to the flow (Figure 30), and coincide

with the walls of the channel. The boundary condition needed to solve the problem is the no-slip condition $u|_{wall} = 0$, which assumes the velocity of the fluid next to the wall is zero.

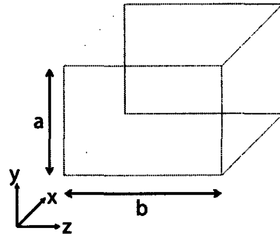


Figure 30. Domain of microchannel simulation. The parameters a and b are the channel wall lengths. Fluid flow is in the x-direction.

To calculate the flow field in the rectangular microchannel, the differential equation is solved by separating the particular and homogenous parts of the solution. A particular solution that works for the channel velocity is found by assuming a parabolic velocity variation in one of the two dimensions (we chose y). We found the homogenous part using the method of separation of variables in the y and z directions. The resulting set of orthogonal solution functions are matched to the boundary conditions and to the particular solution. The solution for fully developed laminar flow in a rectangular channel of arbitrary aspect ratio is:

$$u(y,z) = -\frac{1}{2\mu}\frac{dP}{dx}\left(y^2 - ay\right) + \sum_{k=1}^{\infty} C_k \sin\frac{k\pi}{a}y \cosh\frac{k\pi}{a}\left(z - \frac{b}{2}\right)$$

$$C_k = \frac{2a^2}{\mu k^3 \pi^3}\frac{dP}{dx}\frac{\cos k\pi - 1}{\cosh\frac{k\pi b}{2a}} \quad . \qquad (3.6.2)$$

For the accuracy desired here (< 1% error), we found the first 20 terms in the infinite summation sufficed.

Once the flow problem has been solved, the temperature distribution in the channel can be found. This time instead of performing a force balance on the differential volume one performs an energy balance. The heat flowing along the fluid flow direction is assumed to be entirely due to fluid transport, and not conduction (this assumption should be checked after running the simulation, especially for liquid metals whose thermal conductivity is quite high). In the lateral direction, there is transport from the neighboring fluid elements via thermal conduction. The resulting differential equation is

$$\rho c_P u(y,z)\frac{dT}{dx} = -k\left[\frac{d^2}{dy^2}+\frac{d^2}{dz^2}\right]T(y,z) \tag{3.6.3}$$

Unlike Eq. 3.6.1, the non-homogenous part of this equation is proportional to u(y,z) rather than to a simple constant, making a particular solution problematic. Also, a method of solution allowing the boundary conditions to be easily altered is desirable. We therefore developed a 2-D finite difference solution to the problem by discretizing space in the y and z directions. (A similar finite difference solver for the flow problem was also written, in order to handle more complicated geometries). The boundary condition chosen for the channel was that of constant heat flux into the channel. This is implemented by creating a "fake" boundary one discretization step beyond the true boundary, and setting its temperature such that the input heat flux equals q_s, a predetermined constant.

The temperatures at all points within the channel are calculated using successive approximations. During each step the new temperature in a volume element is expressed in terms of the old temperatures at the surrounding four elements, plus a source term from the non-homogenous part of Eq. 3.6.3. The correct boundary conditions are reestablished at the start of every step. After beginning with a guess (uniform temperature), the simulation runs until the maximum change between successive iterations falls below a threshold. The solution for the temperature is slow, taking around 5 minutes to reach tolerable accuracies. The use of a more sophisticated algorithm could speed convergence considerably.

There are two additional constraints on the solution. Because the heat coming into the system is fixed at q_s, the total heat out of the channel due to fluid flow must be made to exactly match the conduction heat into the system. If it does not, the temperature in the channel will diverge. The other issue is that because we have only used von Neumann boundary conditions to specify the problem, the steady-state temperature of the solution tends to drift off, impeding convergence. This is not a problem with the code; although the shape of the distribution can be determined using the von Neumann boundary conditions, there is an additional undetermined constant that must be determined. One fully specifies the solution by setting the bulk fluid temperature equal to 320 K. The bulk fluid temperature is the temperature obtained when the fluid is adiabatically mixed, that is to say, if the fluid were collected into a cup and stirred before measuring the temperature. We should also note that our simulation assumes all of the heat from the device to be cooled must leave through the water. In a real device, there could also be

other paths (e.g. natural convection or conduction into a larger body) that might aid in heat removal.

To test the simulation, a 100 μm by 100 μm channel was used, discretized with a 150 by 150 grid. Two fluids were tried, water and a liquid metal eutectic, $Ga^{68}In^{20}Sn^{12}$. The relevant thermophysical data of both are shown in Table 2 [60], in SI units.

	ρ	c_p	k	μ
Water	998.0	4181	0.606	9.5808e-4
$Ga^{68}In^{20}Sn^{12}$	6363.2	139.068	39	2.2150e-3

Table 2. Thermophysical data for the two fluids tested.

Assuming a pressure gradient along the channel of 1 bar/cm, and a heat flux of 10 W/cm^2 at the walls, the results shown in Table 3 below are obtained.

	u (m/s)	Re	f_D	Nu	h (W/cm^2/K)
Water	3.67	382	0.149	3.64	2.21
$Ga^{68}In^{20}Sn^{12}$	1.59	456	0.123	3.53	138

Table 3. Channel simulation data for the two fluids tested.

The results indicate that while both water and $Ga^{68}In^{20}Sn^{12}$ do a good job of transporting heat through the channel, the liquid metal offers approximately a 60-fold increase in the local heat transfer coefficient. Since the Nusselt numbers of the two flows are similar, we may attribute the enhancement in the heat transfer coefficient to the higher thermal conductivity of the liquid metal, since $Nu = \frac{hD_h}{k}$.

The numerical accuracy of the computation can be checked by comparing the calculated flow parameters to analytical parameters from a heat transfer text or from similar numerical simulations [67]. Theoretically, the product of the Reynolds number and the Darcy friction factor for the flow in this geometry should be 56.9, and for both fluids the actual calculation yields 56.9. The agreement is excellent, indicating that 20 terms in the exact solution for the flow (Eq. 3.6.2)

are sufficient. The temperature calculations can be checked as well; the theoretical value of the Nusselt number for both flows is 3.61, whereas the calculated values were 3.64 and 3.53. Finer spatial quantization reduces the error. Increasing the discretization mesh density from 50x50 to 100x100 reduces the error of the liquid metal simulation from 8% to 5%, and a further increase of the density to 150x150 reduces the error to 2%.

From the calculated flow parameters shown in Table 3, the hydraulic power needed to drive the flow through a microchannel of length L=1 cm can be computed. Neglecting entrance and exit effects, the pressure drop across the channel is given by

$$\Delta P = \frac{\rho \bar{u}^2 \bar{f}_D}{2D_h} L \ . \qquad (3.6.4)$$

The pumping power W_P can be calculated from the pressure drop, the channel area, and the mean flow according to

$$W_P = ab\bar{u}\Delta P \ . \qquad (3.6.5)$$

For a 1 cm channel, the hydraulic pumping power is calculated to be 1.58 mW for the metal and 3.67 mW for water.

To summarize, we have found that a significant increase in heat transfer coefficient (60x) is expected in a microchannel utilizing the liquid metal eutectic, $Ga^{68}In^{20}Sn^{12}$, in accordance with the results of Miner and Ghoshal[60]. The pumping power required to drive the fluid flow is also lower for the liquid metal, an important consideration for power generation applications. For the remainder of the chapter we will be to using water in our microchannels, but the possibility of using advanced coolants should not be forgotten.

3.7 Modeling an integrated TE generator and microchannel cooler

In this section we apply microchannel cooling and heat spreading to the problem of thin-film thermoelectric power generation. We show that microchannel cooling has the potential both to enable the use of thin-film TE elements and to significantly increase the generated power density of elements whose length is not constrained. Before doing so, we examine the performance of two systems useful as a baseline for comparisons. The first is a commercial system with published specifications, and the second uses the optimization theory we have

developed in the previous sections together with optimal heat spreading and the largest passive heat sink coefficient that we felt was reasonable.

3.7.1 Performance of HZ-2 (baseline commercial system)

Before discussing the design of the microchannel generator, it is worth reviewing the specifications of a commercial thermoelectric generator, Hi-Z's HZ-2. The HZ-2 was shown in Figure 21, and from the specification sheet for the device one can find the information in Table 4.

T_{hot}	508 K
T_{cold}	308 K
Area	8.4 cm^2
Heat spreading area ratio ($A_{TE}/A_{heatsink}$)	0.44
Length	0.50 cm
Heat flux	9.54 W/cm^2
Generated power density	0.30 W/cm^2
Efficiency at max power	4.4%

Table 4. Operational specifications for the Hi-Z HZ-2 thermoelectric generator.

The temperature difference across the Hi-Z generator during their characterization (T_{hot}-T_{cold}) was 200 K. To enable easy comparisons, we will use this temperature drop for all subsequent theoretical designs. We also note that the heat sink that was used to attain the specified performance can be deduced from the parameters above by assuming that the heat sink ambient temperature T_{fluid} is 300 K. We find the thermal resistance of the generator from

$$R_{th,TE} = \frac{\Delta T}{\text{heat flux} \times \text{Area}} = \frac{200\,\text{K}}{80.1\,\text{W}} = 2.50\text{K/W}.$$

The thermal resistance of the heat sink is given by $\frac{1}{h \times \text{Area}}$, where h is the heat transfer coefficient. Using that, we can find the approximate heat transfer coefficient from the equation:

$$\frac{R_{th,TE}}{R_{th,TE}+\frac{1}{h\times \text{Area}}}\left(T_{hot}-T_{fluid}\right)=\left(T_{hot}-T_{cold}\right).$$

One finds that the heat sink used by Hi-Z during their device characterization had a heat transfer coefficient of around 1.2 W/cm^2/K. (Perhaps unsurprisingly, this indicates that the company used a rather aggressive heat sink to characterize their device's performance.)

3.7.2 Performance of generator with "best" passive heat sink (baseline model system)

The commercial system discussed in the previous section provides us with a useful data point, but one might object to it as a baseline for system performance comparisons because it was not necessarily designed with maximum power density as an overriding goal. Efficiency, cost, and other factors may have driven the design to a different regime from the one which interests us.

To best compare with the full microchannel model, we will develop a baseline model system consisting of 18 Bi_2Te_3-based thermoelectric couples (1 couple = 2 elements, one n-type and one p-type) on top of an idealized "best possible" passive heat spreader and sink. A schematic of a section of this system is shown below in Figure 31.

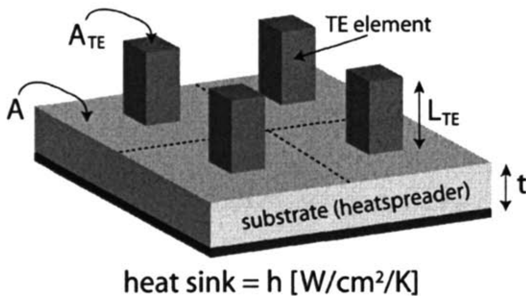


Figure 31. Schematic diagram of TE generator with an idealized heat sink showing important dimensions.

The TE elements are shown as tall rectangular blocks of length L_{TE} and area A_{TE} distributed in a uniform checkered array on a heat spreading substrate of thickness t. Series electrical contacts are assumed but not drawn in. The area beneath each TE element is A, so that the total fractional coverage of the heat sink with thermoelectric material is A_{TE}/A. For simplicity, we will model each TE element with an average magnitude of the material properties of the n and p legs. The material system used throughout this section will be standard bulk Bi_2Te_3, taken from

a study on conventional TE generator performance in the CRC Handbook of Thermoelectrics[68] page 495, and summarized below in Table 5.

	α (μV/K)	κ (W/cm/K)	σ (1/Ω/cm)	ZT (T=300K)
n-type	-190	0.014	741	0.57
p-type	230	0.012	571	0.76
averaged	+/- 210	0.013	645	0.66

Table 5. Thermoelectric material properties for the modeled couples.

The chosen material's thermoelectric properties are intentionally unspectacular compared to those reported from many newer materials and have been chosen to emphasize that the performance enhancement we investigate here is due to the system design rather than the materials. We will take the heat sink parameter h to be equal to 2 $W/cm^2/K$. This is very high relative to the heat transfer numbers typically assumed for natural convection of air or water (see Table 1), and is actually consistent with an optimistic value for pool boiling heat transfer. Nevertheless, we will (somewhat artificially) imagine that we can keep the fluid temperature at 300K. By making this baseline passive heat sink performance very good we will test more stringently whether the active microchannel heat sink can provide added performance.

We begin by modeling a generator consisting of 36 elements, each of them 20 microns thick. This will be the representative thickness for thin-film TE elements in this study. The temperature at the hot side of the thermoelectric is assumed to be 500 K and the temperature that the heat sink is referenced to is 300 K. The heat spreader is chosen to be made of silicon and its thickness is 300 microns. This thickness is chosen because it is thick enough to spread out the heat from the TE elements but small enough that it contributes negligible thermal resistance compared to the heat sink. Because the operating points of interest in this study will be near thermal impedance matched conditions (not strongly heat sink limited) to a good approximation the optimal electrical load resistance for power generation will be given by impedance matching. The unknown parameter to be optimized in this picture is the amount of heat spreading as parametrized by the ratio A_{TE}/A. Figure 32 shows the power density (generated power divided by heat sink area A) and efficiency as a function of the coverage ratio A_{TE}/A.

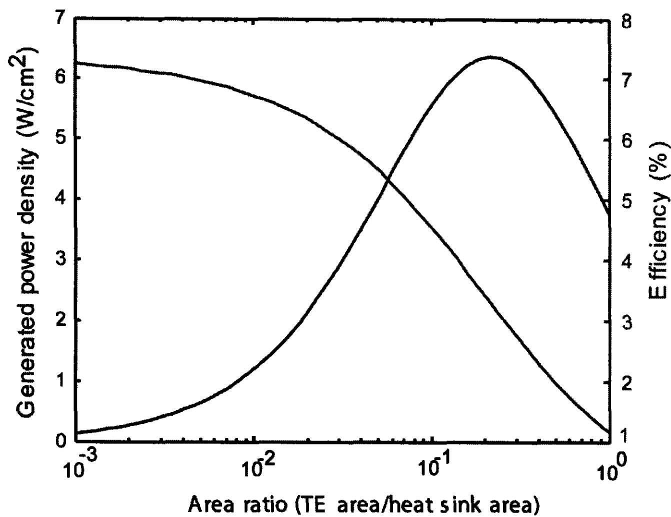


Figure 32. Generated power density and efficiency for benchmark TE thin-film generator, as a function of area coverage.

The maximum power generated in this system is 6.4 W/cm^2, at 3.3% efficiency. The ratio of thermoelectric area to heat sink area (for optimal power generation) is approximately 1/5. With a higher than optimal area ratio (less heat spreading) the heat sink is unable to remove heat from the cold side of the generator fast enough, raising the cold side temperature and degrading power and efficiency. With a lower than optimal area ratio (more heat spreading) the efficiency increases (since the cold side temperature is decreasing) but the small thermoelectric area restricts the heat flow through the generator, degrading the power output. This can also be seen by examining the temperatures of the different parts of the system as a function of area coverage as shown in Figure 33.

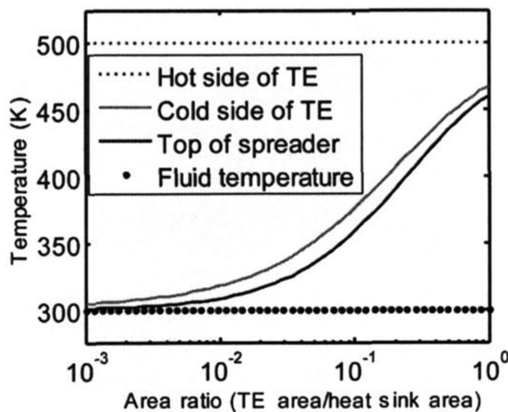


Figure 33. Layer temperatures for a TE generator (benchmark heat sink) as the coverage area of the TE elements on the spreader is varied.

Examining the temperature plot it is clear that at the optimal power point (Area ratio=1/5) the temperature at the cold side of the generator is roughly the average of the fluid temperature and the hot side temperature. In other words the thermal impedance of the thermoelectric elements is roughly equal to the thermal impedance of the combined spreader and sink in accordance with the "thermal impedance matching" prediction of Eq. 2.8.3.

There is one more interesting question to ask before moving on to the microchannel system: what is the maximum power density that we can achieve if we allow the length of the element to vary? That is, if we no longer constrain ourselves to thin-film thermoelectric power generation, how much power can we generate given the constraints of the benchmark heat sink? The answer can be found by setting the area coverage fraction to unity and allowing the length of the TE to vary until the power is optimized (Figure 34). This happens at an element thickness of around 105 microns for this material and heat sink. The power generation density at this element thickness is 7.38 W/cm^2 with 3.75% efficiency. This may be regarded as the best possible power density in the benchmark system. This power density is significantly higher (~25x) than the commercial system mainly due to the thinner element and the better heat sink coefficient.

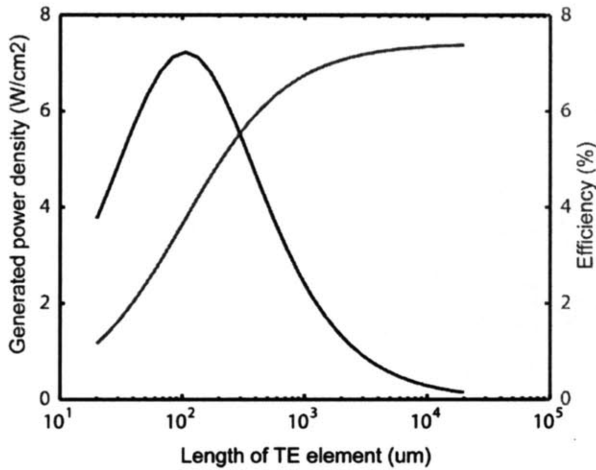


Figure 34. Generated power density and efficiency as the length of the thermoelectric element is varied.

3.7.3 Active microchannel heat sink – geometry and overall model

A schematic diagram of the TE generator, heat spreader, and microchannel cooler we will model is shown in Figure 35.

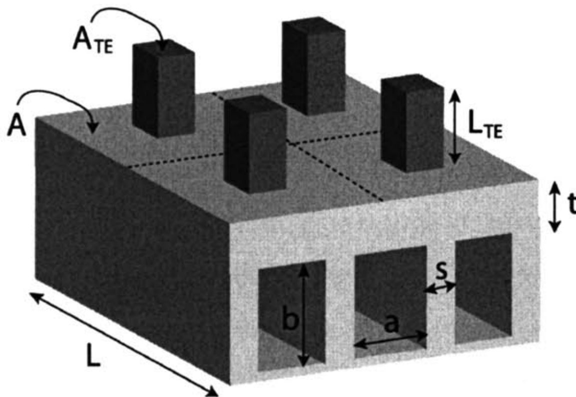


Figure 35. Schematic picture of TE generator integrated with microchannel cooler, showing important dimensions.

The thermoelectric element material is identical to the benchmark material shown discussed in the previous section, and the spacing and size of the elements will be described by the same variable names. Under the thermoelectric elements is a silicon substrate that also functions as a heat spreader (thickness t, κ_{sub}=1.3 W/cm/K) into which microchannels have been etched or

cut. The bottoms of the channels are assumed to be capped but we will later (conservatively) assume that all of the heat transfer is through the tops of the channels or into the channel sidewalls. The water coolant is pumped through n_c=36 parallel channels of width a, height and separation s. The length of the channels is L, and the width of the entire heat sink platform is given by $n_c(a+s)$. In other words the total heat sink width is determined by the channel spacing, the channel thickness, and the number of channels. Once the geometry of the TE elements, the spreader, and the channels is determined, the mean fluid flow velocity through the channels (u) can be set by varying the applied pressure.

The electrical model for the TE elements in Figure 35 is similar to that described in the benchmark; the elements are assumed to be connected electrically series with an impedance matched load in which the generated power is measured. Once again, 36 TE elements are used (18 couples), so that typical closed circuit output voltages are between 0.3 to 0.7 volts (an acceptable output level for a DC-DC converter). There is a crucial new ingredient in the system electrical model: because the microchannels must be pumped we need to keep track of the electrical power that must be spent moving water through the channels (i.e. running the heat sink). There is a certain amount of heat loss which is hydrodynamically required due to frictional losses in the water. This is given by the Darcy friction factor discussed in the previous section. On top of this fundamental hydraulic loss, there are channel entrance and exit losses. Both the Darcy friction factor and the entrance and exit losses are difficult to compute from first principles. The simulation in the previous section can not be relied on for quantitative results both because it assumes a fully developed flow profile (a poor approximation for the microchannel geometries of interest to us) and because we will be interested in the turbulent as well as the laminar flow regime. The detailed modeling of the pressure drop and hydraulic losses are discussed in detail in Appendix 5, and are based on empirical correlations summarized in the literature[48].

The last ingredient in the electrical/pumping model is the "wire-to-water efficiency" capturing the losses in the electrical motor + pump system which is used to force water down the microchannels. The value assumed for this efficiency was 55% based on numbers from high efficiency solar pumps in the literature [69]. We note that for a pump to achieve optimal efficiency the pressure and flow requirements of the microchannel must be typically well-matched with the pump. The pump head and flow specifications for high-efficiency positive displacement pumps (piston pumps, gear pumps) seem to cover the pressures and flows needed for the

microchannels we will consider here. To the extent that mismatches occur we expect that that the number of channels (i.e. the width of the platform) can be varied to improve the match without affecting the results of this section, since we will be concerned with power densities and efficiencies that do not scale with the platform width.

The thermal model for the system shown in Figure 36 is constructed using separate models for each of the thermal resistances between the bottom of the thermoelectric element and the coolant water. Because the TE elements are identical, the problem can be divided up into repeating blocks so that only the area A beneath a single TE element need be considered. We describe each of the thermal resistances in turn by relating the total heat power leaving one thermoelectric element (P_{out}) to the temperatures shown in Figure 36.

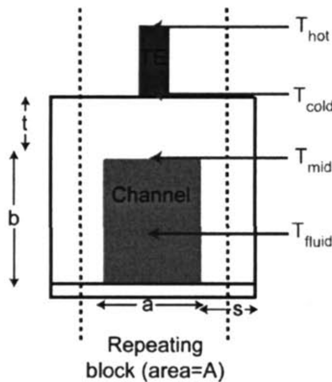


Figure 36. Side view of microchannel cooler showing one of the repeating blocks of the generator (one TE element with its corresponding heat sink area). Also shown are the temperatures used to define the thermal resistances for the heat transport model.

The heat transfer through the thermoelectric elements is assumed to be uniform and one dimensional between the temperatures T_{hot} and T_{cold}. It is due to thermal conductivity, Ohmic heating, and the Peltier effect. It is given by:

$$P_{out} = P_{cond} + P_{Pelt} + P_{Ohmic}$$

$$P_{out} = \frac{\kappa\left(T_{hot}-T_{cold}\right)A_{TE}}{L_{TE}} + \alpha^2\sigma\frac{T_{cold}\left(T_{hot}-T_{cold}\right)A_{TE}}{2L_{TE}} + \frac{1}{8}\alpha^2\sigma\frac{\left(T_{hot}-T_{cold}\right)^2 A_{TE}}{L_{TE}} \quad . \quad (3.7.3.1)$$

Heat exiting the bottom of the thermoelectric element is spread through the substrate, according to the area ratio A_{TE}/A. The spreading contributes an additive thermal constriction resistance that we model with the steady state spreader resistance $R_{th,spread}$ given by Eqs. 3.2.2 and 3.2.3, so that

$$P_{out} = \left(T_{cold} - T_{mid}\right)/R_{th,spread} \quad (3.7.3.2)$$

After the spreader the heat can pass into the water either directly through the top of the microchannel or through the silicon fins dividing the channels and then through the channel sidewalls. The joint process is modeled with an effective heat transfer coefficient h_{eff} defined so that the total heat transfer below one TE element into the microchannels is given by:

$$Q_{out} = h_{eff} A\left(T_{mid} - T_{fluid}\right) \quad (3.7.3.3)$$

T_{mid} refers to the mean temperature in the substrate at the plane that is level with the top of the channel. T_{fluid} denotes the average fluid temperature down the length of a channel. It is not a priori obvious that it is reasonable to use a lumped channel temperature to approximate heat transport into the water but it will turn out that the water temperature is fairly constant compared to the other temperatures in the system. (For the same reason we do not model any changes in the thermophysical parameters of water resulting from the temperature change in the fluid.) The effective heat transfer coefficient (h_{eff}) defined in this manner is not to be confused with the heat transfer coefficient from the silicon into the water at the walls of the microchannel (h). As discussed in the previous section, microchannels enhance both the heat transfer into the water (resulting in a large h) and the effective area over which that heat transfer occurs (resulting in an even larger h_{eff}). Following the microchannel and heat sink literature[55,70] we model the effective area "seen" by the heat at the bottom of the spreader as

$$A_{eff} = (2\eta_{fin} b + a)Ln_c, \quad (3.7.3.4)$$

so that $h_{eff}A=hA_{eff}$ (actually our definition of h_{eff}). Here the "fin efficiency" parameter η_{fin} is used to capture the imperfect heat transfer though the silicon fins dividing the channels; if the thermal conductivity of the substrate material was infinite the fin efficiency would be equal to 1. Its value is given by

$$\eta_{fin} \approx \frac{\tanh\left(b\sqrt{2h/s\kappa_{sub}}\right)}{b\sqrt{2h/s\kappa_{sub}}}, \quad (3.7.3.5)$$

where h is the heat transfer coefficient into the water from the channel walls and κ_{sub} is the thermal conductivity of the substrate. For small values of the term $b\sqrt{2h/s\kappa_{sub}}$ the fin efficiency is close to one, at $b\sqrt{2h/s\kappa_{sub}} = 2$ the fin efficiency is about 50%, and for larger values it drops quickly to zero. This can be understood physically. If the fin thermal conductivity or the fin

thickness is increased, the heat can travel further down the fin before being absorbed into the adjacent microchannel, which is equivalent to an increase in fin efficiency. Similarly, if the heat transfer coefficient into the channel is increased, the heat will not travel as far, decreasing fin efficiency. (The dependence of the fin efficiency on b is just due to its definition in Eq. 3.7.3.4 as the fraction of the channel height which is seen by the heat.)

The next ingredient needed in the description of the thermal transport is the heat transfer coefficient into the water, h. This is given by Nusselt number correlations from literature, according to the relation:

$$h = \frac{\kappa_{water}\, Nu}{D_h} = \left[\kappa_{water}\, Nu\right] \Big/ \left[\frac{4ab}{2(a+b)}\right]. \qquad (3.7.3.6)$$

Care must be taken to account for the change in Nusselt number that occurs between the laminar flow and turbulent regimes (Re>2300), and for the variation of Nusselt number with developing flow. The detailed Nusselt number correlations were again taken from Incorpera[48], and are discussed in detail in Appendix 5. The final piece of the heat transfer model is the equation describing the heat as it exits the channels through the water and the corresponding rise in the mean temperature T_{fluid} of the water:

$$\begin{aligned} Q_{out} &= n_c\left(T_{out} - T_{in}\right)\dot{m}c_p \\ T_{fluid} &= \left(T_{out} + T_{in}\right)/2 \end{aligned} \qquad (3.7.3.7)$$

Here $\dot{m}$ refers to the mass flow of water through one channel [kg/s] and c_p is the specific heat of water [J/kg/K]. These equations are included to keep track of the temperature rise in the coolant along the length of the channels. The inlet temperature is set to 300 K for all simulations.

We turn now to the simulation results for the coupled generator and microchannel heat sink. While a formal and exhaustive optimization of the model (simultaneously varying all independent design parameters to maximize power) was not attempted, by varying a few carefully chosen model parameters we can design a pretty good generator. We found that for a 36 element generator made from the L_{TE}=20 μm thick Bi_2Te_3 material described in Table 5 of the previous section the following design worked well: a=300 μm, b=1000 μm, s=100 μm, t=100 μm, u=6.5 m/s, A_{TE}=A (no heat spreading), and L=1 cm. A hot side temperature of 500 K and a fluid inlet temperature of 300 K were assumed.

The maximum net generated power density is 37.9 W/cm^2, and the efficiency at that power was 3.3%. The total generated power was 49.5 W/cm^2, but 11.6 W/cm^2 was spent pumping the channels. We note that the outlet temperature of the water rises only about 5 degrees centigrade, justifying our assumptions regarding the averaged heat transport into the channel. (We'll look more at the performance of the generator at and near this design and operation point in the next few sections.)

The net generated power of the microchannel generator is more than 6 times the power generated in the passive, idealized benchmark model of given in Section 3.7.2. It is more than 120 times the power density generated by the Hi-Z commercial system discussed in Section 3.7.1.

We can better understand the operation of the system we have designed by studying how the performance changes around the design point as certain key parameters are varied.

3.7.4 Active microchannel heat sink – flow optimization

As the mean flow velocity through the channels is increased, two things happen. The Nusselt number associated with heat transport into the channel also increases, increasing the temperature drop across the thermoelectric elements, and generating more power. However, the amount of power needed to pump the channels increases with increasing velocity roughly as the square of the flow velocity, eventually overcoming the benefits of the enhanced heat sinking. Figure 37 shows the performance as a function of the flow velocity u.

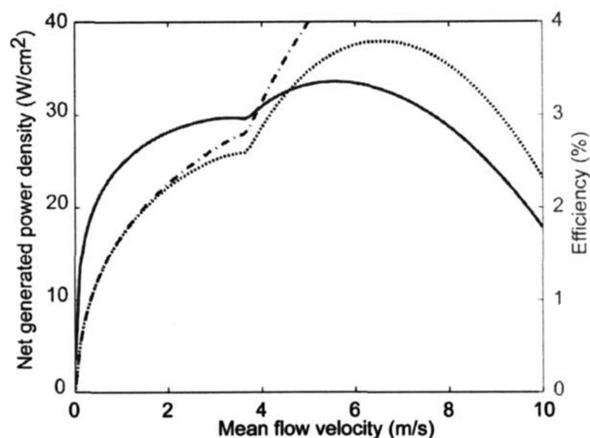


Figure 37. Efficiency (solid line), generated power (dash-dot line), and net generated power density (dotted line) for varying mean water flow velocities. Net generated power is defined as generated power minus the power needed to pump the water through the channel.

Both the efficiency (solid line) and the net generated power density (dotted line) display clear maxima, as we expected. The kink in the curved near u=4 m/s is due to the onset of turbulent flow in our model. The key physics relevant for the flow velocity is most evident in the interplay between three dimensionless parameters: the Nusselt number (describing the enhanced heat transfer into the channel), the Darcy friction factor (describing the power required to pump the channel), and the Reynolds number which scales with the flow velocity. They are plotted below in Figure 38. Of particular interest is the effects of the onset of turbulence at Re=2300, leading to an abrupt increase in the Nusselt number. Note also that at high Reynolds numbers (highly turbulent flow), the Darcy friction factor decreases much more slowly (for laminar flow $f_D \sim 1/Re$), essentially leveling off. This is because the turbulent flow is effectively isolated from the shear force exerted by the walls. For normal laminar flow, the mean flow velocity increases linearly with the applied pressure gradient, a hydraulic equivalent of Ohm's law for resistors. For turbulent flow this relationship breaks down, and the velocity increases more slowly with the pressure gradient. Equivalently, for an increasing velocity the pressure gradient grows more rapidly (almost quadratically for fully turbulent flow). This means that the pump power (computed from Eq. 3.6.4 and 3.6.5 and scaling as the pressure drop times the flow velocity) increases almost with the cube of the flow velocity, which quickly becomes energetically costly. We should also note that the Darcy friction factor in real measurements often displays erratic and discontinuous behavior near the cross over to turbulence. The correlation we have used intentionally hides this issue by smoothly transitioning between flows (Appendix 5). While this is not necessarily physical, it makes it easier to rationally optimize the generator in the difficult-to-model flow transition regime.

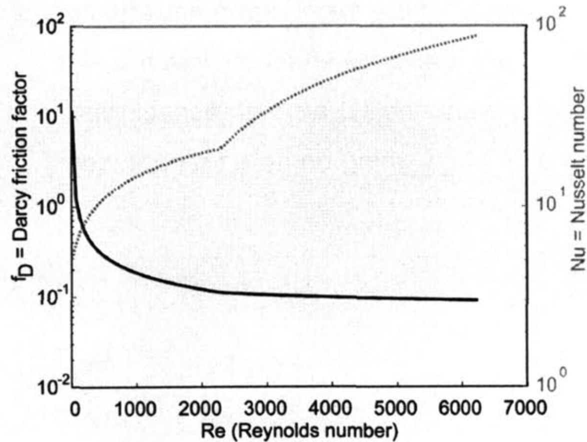


Figure 38. Nusselt number and Darcy friction factor as a function of Reynolds number. Note the effects of turbulence, modeled here at Re>2300.

3.7.5 Active microchannel heat sink – channel aspect ratio optimization

Another interesting parameter to vary is the aspect ratio of the microchannels (b/a) while keeping the total cross-sectional area of the microchannels fixed (ba=constant). An increase in the aspect ratio creates longer fins, increasing the effective area seen by the heat coming from the TE elements and hence increasing the effective heat transfer coefficient h_{eff} up to a point. However, extremely high aspect ratios also require large pressure drops to achieve the same flow velocity, increasing the pumping power. The result is that there is an optimum aspect ratio for both efficiency and net generated power as shown in Figure 39. The optimum aspect ratio for net power density is evidently near b/a=3.

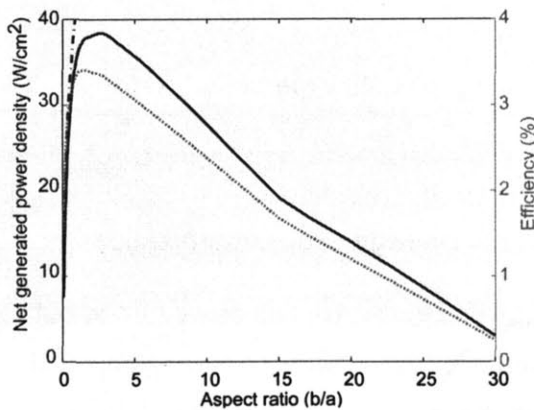


Figure 39. Efficiency (solid line), generated power density (dash-dot line), and net generated power density (dotted line) for various microchannel aspect ratios. Flow velocity is held constant.

We can get a closer look at how the heat sinking performance of the microchannel is affected by aspect ratio by plotting how the channel heat sink coefficient h and the effective heat transfer coefficient h_{eff} (which includes the area enhancement of the channels) vary with aspect ratio. This is shown below in (Figure 40). There are a couple of things going on here that are worth pointing out.

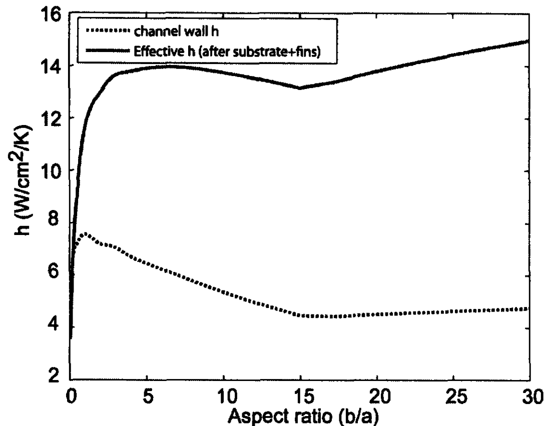


Figure 40. The effects of varying aspect ratio on the channel heat sink parameter h and on the effective heat sink parameter heff, which includes the area enhancement from the fins.

As expected, the effective heat transfer coefficient increases rapidly with increasing aspect ratio until the finite thermal conductivity of the silicon fins limit its effectiveness and it levels off. For reasons that are not physically obvious, the heat transfer coefficient Nu for the turbulent flow which prevails in this simulation is weakly maximized for channels of aspect ratio near unity, which is the reason for the peaked behavior of the channel heat transfer near unity aspect ratio. At large enough aspect ratio (b/a>15) the Reynolds number drops below the threshold for turbulence, and laminar flow is restored.

3.7.6 Active microchannel heat sink – heat spreading (coverage area) optimization

Surprisingly, the optimal heat spreading for the 20 μm thick Bi_2Te_3 used in this study turned out to be none at all. In other words, for optimum performance, we found that the thermoelectric element area A_{TE} should cover (if possible) the entire area of the heat sink A beneath it. This is shown in Figure 41. This is a consequence of the large effective heat sink parameter (h_{eff}~15) seen by the thermoelectric elements. The thermal impedance matching condition is met for this length element if $h_{eff} \sim \kappa/L_{TE}$=6.5 W/cm^2/K. because the effective heat sink parameter exceeds this number, no heat sinking is required to thermally impedance match the heat sink to the

generator. If an even shorter element length is used (L_{TE}=5 μm), then this will not be true, and some heat spreading will be advantageous (Figure 42).

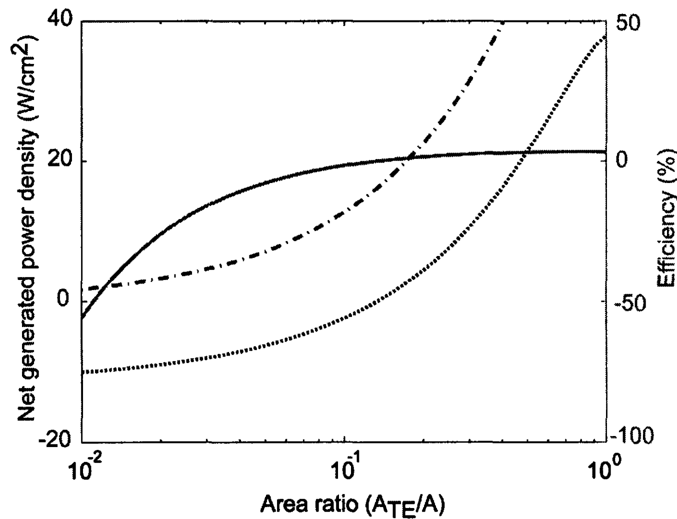


Figure 41. Performance of the microchannel cooled thermoelectric system as a function of the amount of heat spreading used. The optimum net power density is at an area ratio of unity, meaning that heat spreading is not helpful here.

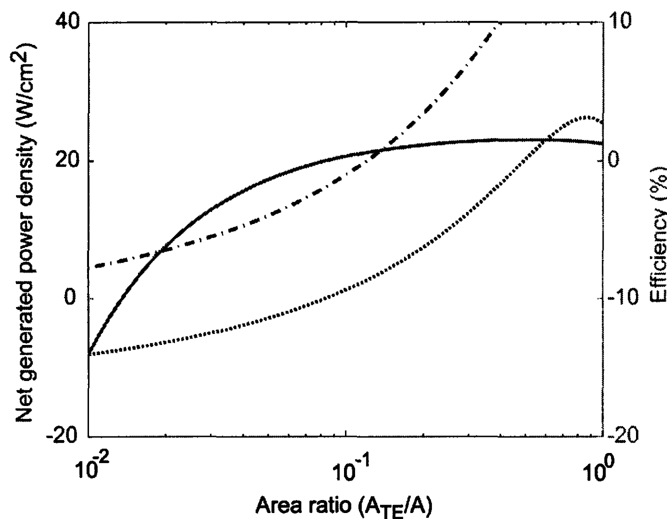


Figure 42. For a shorter (5 micron thick rather than 20 micron thick) thermoelectric element, heat spreading may make sense.

3.7.7 Active microchannel heat sink – thermoelectric element length variation

In discussing the performance of the microchannel cooled TE generator up until now, we have fixed the length of the thermoelectrics at 20 μm. The idea behind this choice was partly that for

many new thin-film thermoelectric materials, thicker elements might not be practical. It is interesting to relax this constraint and study at how the performance of the generator varies with changing element length. This is shown below in Figure 43.

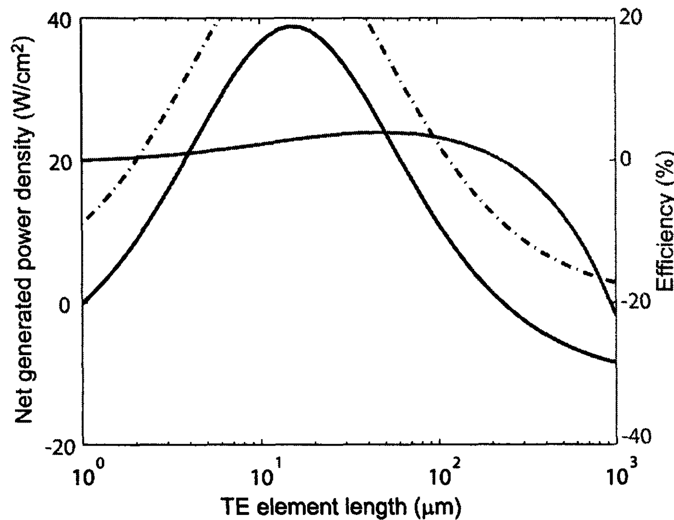


Figure 43. The generated power density and efficiency of the thermoelectric microchannel generator as the length of the thermoelectric elements are varied. The efficiency becomes negative at long element lengths due to the fixed power lost pumping the microchannels.

The power density has a clear optimum between 10 μm and 20 μm. This is actually consistent with our earlier observation that no heat spreading was needed (or desired) to optimize performance of the 20 μm thick generator with the microchannel cooler design chosen here. The efficiency has a higher optimum, which is not surprising in light of the discussion of element length optimization in Chapter 2. There is one key difference with the idealized passive heat sink generator considered earlier in Section 3.7.2. The efficiency in the passively heat sunk generator monotonically increased with increasing element thickness, approaching the limit set by the ZT of the material (see Figure 34). In the current actively heat sunk microchannel generator, as the length of the elements increases, the finite pumping power needed to run the channels actually consumes more power than the thermoelectric is capable of generating, driving the efficiency negative. The difference between the two situations can be highlighted by plotting the power density vs. efficiency loop curves for both situations, as shown in Figure 44.

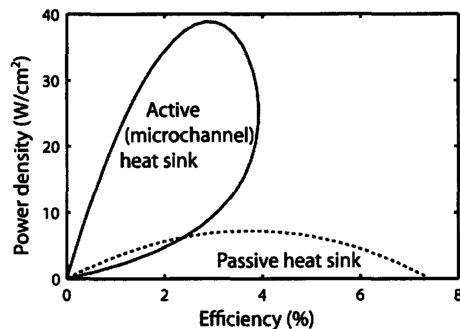


Figure 44. Power density vs. efficiency, with the variation of TE element length used as a parameter. The dotted curve is the TE generator with the ideal passive heat sink, and the solid curve is the microchannel actively cooled generator.

In the case of the actively cooled microchannel generator the efficiency vs. power loop is "closed" at the origin, while the passive generator efficiency curve terminates at a finite efficiency. The reason for the fundamentally different shapes of the loops can be related to our previous discussion of the power vs. efficiency curves for the gas turbine Brayton cycle shown in Figure 14 in Section 2.3. Recalling that discussion, the imperfect isentropic efficiency of the turbine generator was responsible for the "closing" of the curves. The source of that imperfect isentropic efficiency in that model was a "finite heat leak," essentially a shunt path through which heat could move through the turbine at a finite rate without doing useful work. That is also a role played by the thermal resistance of a thermoelectric module. In the case of the passive heat sunk generator, as the length of the elements is increased the heat leaking through the elements decreases, so that even as the generated power drops the efficiency remains finite. In the actively pumped microchannel we have introduced a new isentropic inefficiency (the power required to pump the channels) so the curves do terminate at the origin.

3.8 Comparisons with the literature

Although ours is the first analysis of a coupled thermoelectric generator and microchannel cooler that we are aware of, we could test the microchannel simulation theory developed in Section 3.7 has been against two results from the microchannel literature. We begin by examining the original seminal experimental results of Tuckerman and Pease[62] and then examine a more recent paper in this area[71]. Both papers deal with channel geometries, flow conditions, and materials that could be relevant for the application to thermoelectric power generation. We conclude that there is good agreement between the heat transfer predictions of our theory and the literature examined here as long as the thermophysical property variation of

the water along the channel is not too severe. (Fortunately this was not the case for any of the simulations of the previous section.) This is important because the generated power and efficiency are both very sensitive to the calculated heat transfer coefficient of the microchannels. The agreement between our predicted pressure drop and the experimental pressure drops in the literature were less good (~ 30% difference between the predicted vs. measured friction factors). The reason for this is unknown, although some possibilities are discussed below. It is worth pointing out that the error is such that we appear to be overestimating the pressure drop required to sustain a given flow rate. If that is true we are also overestimating the pumping power required to maintain channel flow in the previous section, which would increase both the efficiency and the generated power of our design slightly.

Tuckerman and Pease (1981)

Tuckerman and Pease demonstrated the first microchannel heat sink in 1981. They fabricated one hundred 1 cm long microchannels on a 1 cm^2 silicon wafer. One set of channels were 302 μm in depth, 50 μm in width, and spaced 50 μm from channel to channel. The top of the channels was 100 μm from the top of the silicon wafer. They applied pressures of up to 31 psi (just over 2 atmospheres) of pressure to the channels and a 790 W heat load via an on-chip resistive heater. At the highest pressure and flow rate they measured a 0.090 K/W thermal resistance between the hottest part of the chip surface (near the channel outlets) and the inlet water temperature. The geometry and flow regime of their work is quite straightforward to model using the tools developed in Section 3.7. Figure 45 shows the original Figure 2 from the Tuckerman and Pease paper, with the results of our simulations overlaid (dotted line).

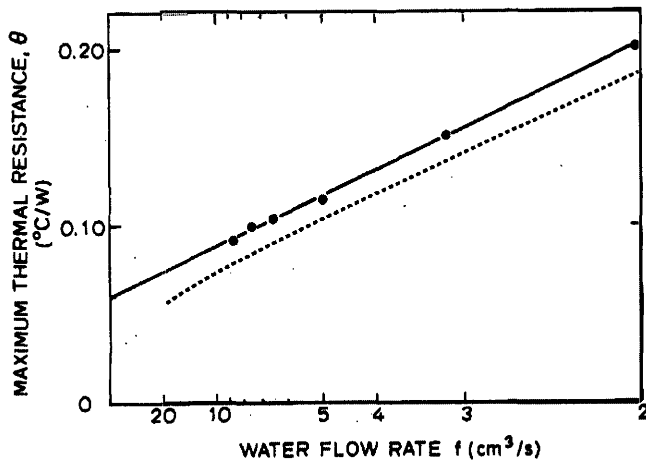


Figure 45. Comparison between the measurements of Tuckerman, Pease (1981) and the microchannel model developed for this thesis. The maximum thermal resistance between the surface of the microchannel heat sink and the inlet water is plotted as a function of channel water flow rate.

The large dots are the Tuckerman and Pease measured thermal resistances, and the solid line is their linear fit to those measurements. The main result of interest is the extraordinarily low thermal resistance they have measured (0.090 K/W at 8.6 cm^3/s flow rate). They also argue that the linearity of the measured data justify treating the channel hydrodynamics as ideal Poisseuille flow (linear pressure-flow relationship) as follows: They first claim that neither the thermal resistance of the silicon nor the Nusselt number characterizing fully-developed heat transfer into the channel should significantly change with flow rate. Since the remaining component of the total heat sink thermal resistance (the thermal resistance of presented by the water itself) decreases linearly with increasing flow rate they plot their observed thermal resistances against the 1/(flow rate). They observe the desired straight line relationship and suggest that their straight line is evidence of ideal Poisseuille flow.

The dotted line is a plot of our theoretically predicted thermal resistance versus the "true" (temperature dependent, channel length dependent, channel aspect ratio dependent) flow rate in the channel. The line under-predicts the thermal resistance of the heat sink by around 0.02 K/W, but appears to have the correct slope.

There are a few interesting points regarding the interpretation of the results, and some open questions we have found in comparing the results of our simulations to the measurements of

Tuckerman and Pease. We found in our simulations that the flow considered by Tuckerman and Pease is not really fully developed. As a result, both the local Nusselt number and the dimensionless friction product f_DRe change substantially through the channel (~30%), especially at higher flow rates. However, these effects are not very apparent in Figure 45 because the principal contribution to the total thermal resistance is from the water between the outlet and inlet port (Figure 46).

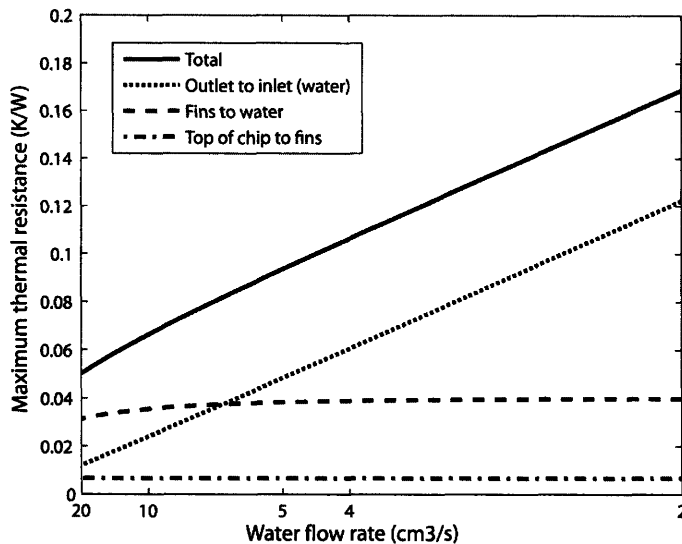


Figure 46. Breakdown of the thermal resistances of our theoretical calculation of the Tuckerman and Pease thermal resistance.

Tuckerman and Pease present some simple theory relating the flow rate to the pressure drop across the channels using the infinite parallel plate approximation for laminar flow and using a constant (temperature independent) viscosity for water for modeling purposes. Later on, they state succinctly during their analysis that "We confirmed the flow rate obeyed Poiseuille's equation and that the thermal resistance was independent of power level." This is surprising for two reasons. First, for the aspect ratio they actually used (channel width/channel height=1/6), using the ideal parallel plate approximation one would expect around a 20% error in the slope of the pressure vs. flow rate relationship[72] in addition to any error caused by the pressure drop at the inlet and outlet ports and channel bends. Second, for the heat flux (790 W) and thermal resistance range they examined in their data, they observed temperature differences of up to 158 degrees C between the top of the heat sink and the fluid inlet temperature. Since they calculate that roughly ¼ of that temperature difference occurs along the length of the channel, presumably the temperature change in the water along the channel length could be as high as

40 C. As shown below in Figure 47, the thermophysical properties of water display significant variation for such temperature differences. In particular, the dynamic viscosity varies by more than 30%. According to our simulations this should have significantly altered the ideal linear Poiseuille relationship between pressure and flow that Tuckerman and Pease observed. This variation is accounted for in an average sense in our simulations because we use the thermophysical properties of water at the mean fluid temperature to calculate all of the flow and heat transfer.

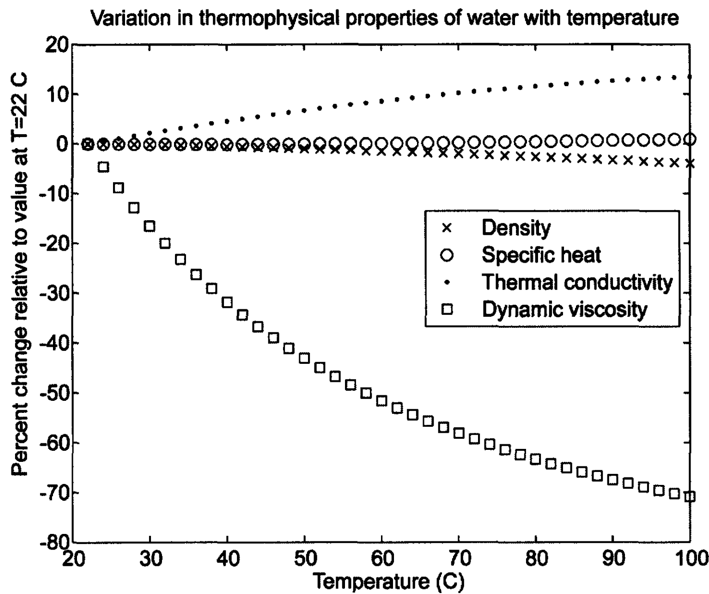


Figure 47. The percent change in the thermophysical properties of water from their room temperature values as temperature is varied.

Another significant difference between the Tuckerman and Pease results and our theory was the pressure drop they reported. They measured a pressure drop of 31 psi at a flow rate of 8.6 cm^3/s. Our simulated pressure drop (corresponding to a flow rate of 8.6 cm^3/s) was 49 psi. Even if all of the non-ideal hydraulic resistance (bending loss, input and outlet loss) are removed in the calculation of pressure drop the theoretical pressure drop we compute (40 psi) is still higher than their measured result. (It is interesting that if the pressure drop is computed using the ideal (infinite parallel plate) Poiseuille flow assumption the pressure drop is 31 psi, exactly what they have measured.) There are two reasonable explanations for this. First, it is also important to realize that for a fixed pressure difference flow rate in the laminar regime is proportional to pipe diameter raised to the fourth power. A small error in the pipe diameters reported by Tuckerman and Pease (50 μm x 50 μm) therefore results in a large change in the

pressure for a given flow rate. Second, there might simply be an error in the friction factor correlation used to find the pressure drop. While the fully developed flow friction factor is likely correct, the channels are short enough that the friction in the developing section of the channel makes a substantial contribution to the pressure drop. In any case, we are not too concerned with the mismatch between the pressure drops for this system, and the heat transfer predictions are close enough for the purposes of our modeling.

Qu and Mudawar (2002)

An experimental study on the pressure drop and heat transfer in a single-phase microchannel heat sink was presented by Qu and Mudawar in 2002. They also present a three dimensional finite difference calculation of the temperature field in the solid and the liquid, to try and include the effects of local heat transfer in the description of the flow and heat transfer. They considered 21 channels, each with a height of 713 μm and a width of 231 μm, so that the width of the cooling platform was 1 cm. The channels were spaced by 236 μm, and their length was 44.76 mm. The inlet water temperature was 15 C. They tested two input heat flux levels (100 W/cm^2 and 200 W/cm^2), and measured the temperature both at the inlet, the outlet, and in the copper substrate at a distance of 2462 μm above the top of the channel for a variety of flow rates. We will compare our simulated results for the 100 W/cm^2 power input with their measured results.

Figure 48 shows a figure from Qu and Mudawar's paper (Figure 4b) with their experimental (Exp) and theoretical (Num) predictions for the pressure drop across the channels as a function of the flow rate, as parametrized by the Reynolds number. Once again, our numerical code seems to overestimate the pressure drop required to drive a given flow rate, especially at high Reynolds numbers. This is true for both the calculated pressure drop that includes an estimate of bend, entrance and exit losses (solid black line) and the ideal pressure drop that includes only the hydraulic pressure drop in the channels (dash-dot line). There are a few possible reasons for this. First, we have used the same numbers for the parasitic pressure drops at the inlet, outlet, and bends that we use everywhere else (see Appendix 5). These do not necessarily exactly match the geometry of Qu and Mudawar, or their theoretically assumed values (which they do not report). Second, the deviation in our theory and their experiment largely disappears if we assume fully developed laminar flow. For higher Reynolds numbers the developing region of the channel is a larger fraction of the total channel. If the correlation for the

developing flow friction factor were too high it would explain the deviation we observe. This is actually the same observation that we made for the Tuckerman and Pease results discussed above. While we will stick with our published correlation for the friction factor in the developing region[72], for future modeling work this regime may be worth reevaluating since even today there is a surprising amount of controversy[73]. Another interesting effect that factors in to the pressure drop is the change in viscosity of water with temperature. At higher flow rates the temperature increase in the 15 C inlet water is less so that the viscosity is higher and the pressure drop greater. It turns out that a 10 degree C error in the fluid temperature is sufficient to explain the slope error we observe in the Pressure vs. Reynolds plots. With the large temperature changes observed in the channel for the chosen heat and flow regime of Qu and Mudawar (up to 40 C temperature change between the inlet and outlet), such an error might also explain some of the error. The subject of how to model channel flow with thermophysical property variation is also still an active area of research[74].

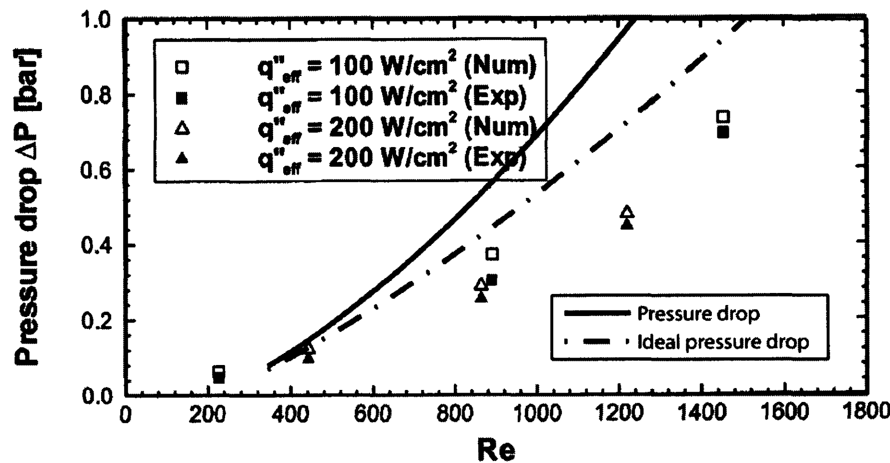


Figure 48. Comparison of microchannel simulation code developed for this thesis and Figure 4 (b) from Qu and Mudawar (2002). The symbols correspond to Qu and Mudawar's measurements and also to their theoretical predictions of the pressure drop across the microchannels for two different input heat powers (100 W/cm^2 and 200 W/cm^2). The solid line depicts the prediction of out theory for 100 W/cm^2 of input power.

Figure 49 (corresponding to Figure 5 in the paper of Qu and Mudawar) plots the change in water temperature between the inlet and the outlet as a function of Reynolds number (flow rate). The dotted line is the theoretical prediction of our theory for 100 W/cm^2 of input heat flux. There is excellent agreement at the higher Reynolds numbers but the curves begin to diverge for low Reynolds numbers. At low Reynolds numbers the flow rate is low and the water spends more time in the channel so that the temperature difference between the inlet and outlet is larger. We

believe the error is due to the variation in the viscosity of the water with the temperature along the channel which is not accurately captured by our simplified averaged self-consistent thermophysical property method.

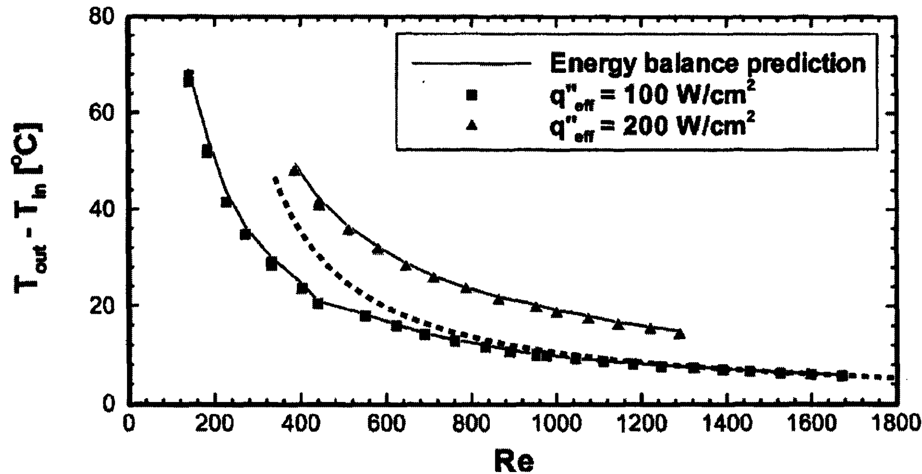


Figure 49. A comparison of the microchannel simulation code developed for this thesis to the results of Qu, Mudawar (2002). The plot depicts the total temperature change in the water flowing through the microchannels betwen the inlet (T_{in}) and outlet (T_{out}). The symbols (measurements) and solid lines (theory) are from Qu and Mudawar, and correspond to two different levels of input heat flux (100 W/cm^2 and 200 W/cm^2). The dotted line is our theoretical prediction corresponding to an input heat flux of 100 W/cm^2.

Figure 50 (Figure 6a from the Qu and Mudawar paper) depicts the experimental temperature measurements (called T_{th}) of Qu and Mudawar from four thermocouple probes positioned in the substrate 2.462 mm above the top of the microchannels relative to the inlet fluid temperature T_{in}. The measurement was for an input heat flux of 100 W/cm^2. Location 1 corresponds to a point 5 mm from the inlet port and locations 2, 3, and 4 are spaced 11.588 mm from one another along the channel. Our theory seems to do a reasonably good job predicting the average temperature rise of the surface of the microchannel cooler except for at low Reynolds numbers. This deviation is actually due to the one anomalous inlet and outlet temperature rise discussed in Figure 49.

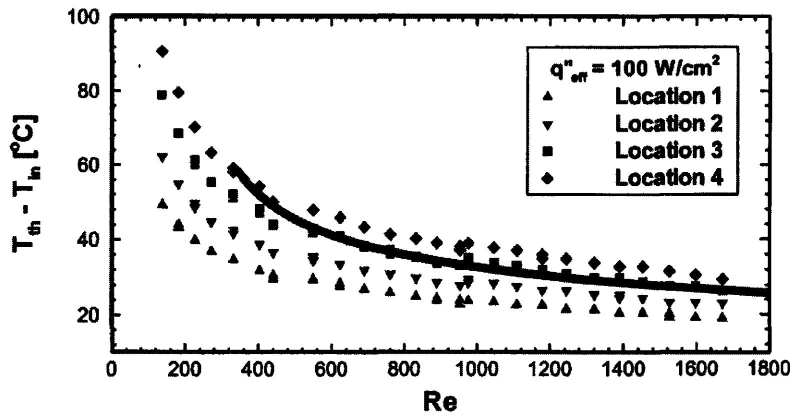


Figure 50. A comparison between the microchannel heat sink theory developed for this thesis and the results of Qu and Mudawar (2002). The symbols depict the experimental temperature measurements (T_{th}) of Qu and Mudawar from four thermocouple probes positioned in the substrate 2.462 mm above the top of the microchannels relative to the inlet fluid temperature T_{in}). The measurement was for an input heat flux of 100 W/cm^2. Location 1 corresponds to a point 5 mm from the inlet port, and locations 2, 3, and 4 are spaced 11.588 mm from one another along the channel. The solid line shows our theoretical prediction for the average temperature difference between the top of the microchannel and the inlet fluid temperature.

3.9 Conclusions

In this chapter, the key obstacle to thin-film thermoelectric power generation (and high power density thermoelectric generation in general) was identified as insufficient heat transfer out of the cold side of generator. A similar, but usually less severe problem exists for heat transfer into the generator. Two approaches were suggested for overcoming this difficulty: heat spreading and microchannel heat sinking.

Analytic models for heat spreading were introduced, both for the steady-state regime and for more general dynamic problems. Heat spreading as a means of thermally impedance matching a thin-film to a given heat transfer coefficient was investigated with a model of a single TE element and spreader. Our model demonstrated that substantial gains in generated power and efficiency are sometimes possible using a well-designed spreader.

Improvements to the heat transfer coefficient of the heat sink were investigated as a means for improving thermoelectric power generation. A method for the experimental determination of the heat transfer coefficient was presented along with some experimental results. Microchannels were advanced as a solution to the extreme heat transfer problems posed by thin-film

generators. The fluid flow and heat flow in a laminar microchannel were modeled and the model was used to investigate the applicability of liquid metal as a coolant. It was found that the thermophysical properties of the metal eutectic can substantially enhance the heat transfer into the fluid from the wall of the microchannel while decreasing the required pumping power.

Lastly, both the heat spreader and the microchannel concepts were combined with the thermoelectric element model of the previous chapter to model a full generator with optimized heat spreading into a microchannel heat sink. Empirical correlations were used to model the channel flow in both the turbulent and laminar flow regimes and lumped models were used for the three dimensional thermal resistances. This strategy (as opposed to a 3D finite element or more complete computational fluid model) enabled the design parameter space to be examined quickly, while retaining a grasp of the basic physics behind the results. The microchannel model heat transfer predictions (most important for the thermoelectric generator) agreed reasonably well with experimental results from the literature while the predicted pressure drops seemed to err somewhat (albeit on the conservative side from the standpoint of power generation). Some of this error may have been due to significant thermophysical property variation with temperature (not present in the microchannel TE design) and due to uncertainties in the measured channel diameters in the microchannel literature.

The performance of the actively pumped microchannel generator was compared with an optimistic passive heat sink design, and a greater than 6-fold increase in net generated power density was demonstrated. Compared to a commercial generator operating over an identical temperature drop and made from similar material, the power generation density increased by over a factor of 120. Also, despite the power used to pump the channels, the overall efficiency was still 3.3%, comparable to both the idealized passive generator (3.75%) and the efficiency claimed by the commercial module (4.5%). It would appear that the microchannel generator design substantially increases generated power density while not severely degrading the efficiency. Ironically, this is due to the fact that thermoelectric generators are relatively inefficient generators to begin with (compared with, for example, the gas turbine Brayton cycle discussed in Chapter 2). The resulting ease with which small amounts of efficiency can be traded for large gains in power density offers additional support for our claim that power density is often a more rational optimization goal than efficiency for thermoelectric generators.

Chapter 4: Thin-film TE performance characterization

4.1 Overview

This chapter presents measurements made using instruments we have built to perform real conditions testing on thermoelectric elements and generators. In Section 4.2 we review some of the methods used for thermoelectric material characterization, and discuss some of the issues with the application of these methods to thin films. Section 4.3 introduces the tools used for the measurements, along with some key elements of their design and operation. Section 4.4 presents measurements on bulk thermoelectric elements obtained from commercial sources. The dependence of generated power on length scaling is investigated. The high temperature Seebeck coefficient and thermal conductivity of Marlow's micro-alloy material are measured and compared with Marlow's low temperature characterization data. Individual elements removed from a commercial thermoelectric generator are characterized and used with a model of the generator in order to predict the generator performance. Section 4.5 discusses the "thick-film" measurements made on lead-salt based thermoelectric materials from the Harman group at Lincoln Labs. Section 4.6 discusses thin-film measurements made on ErAs nanodot epitaxial layers on InP substrates, and on a thin-film generator made from this material.

4.2 Thermoelectric characterization techniques

One of the main challenges faced while developing new thermoelectric materials is the accurate and repeatable measurement of the material characteristics $\alpha(T)$, $\sigma(T)$, and $\kappa(T)$. Without a trustworthy standard for these parameters, optimization of growth conditions for thermoelectric performance is not possible. For macroscopic samples (>1 mm in length) there exist excellent methods for the measurement of these parameters. For microscale samples (e.g. thin-films) there appear to be few measurement techniques accepted as a standard in the field. Measurements of microscale thermoelectrics are fundamentally more difficult and there are several serious issues that arise when attempting to apply macro-scale measurements to micro-scale samples.

Transient Harman technique

The standard characterization method for macroscopic thermoelectric elements in both the commercial and academic worlds is the so called Harman technique[75]. It is exceptional because it measures the thermoelectric figure of merit Z directly, and does so using only electrical voltage measurements. It was originally advanced as a specialized method for thermal conductivity measurement in thermoelectric materials, the idea being that the other parameters in ZT (σ, α, T) could be easily measured. We provide a simplified discussion of the technique in order to understand the problems encountered in extending its use to thin films.

An alternating current I excites the element at a variable frequency, with small enough amplitude so that the perturbation of the current on the dc (mean) temperature of the sample can be neglected. If a dc (or very low frequency) current excitation is applied to the device, the amplitude of the voltage across the device will be given by the sum of two voltages, the Ohmic voltage drop due to the element's finite resistance R, and the Seebeck voltage due to the temperature difference ΔT which builds up across the device due to the thermoelectric transport.

$$V_{slow} = V_{Ohmic} + V_{Seebeck} = IR + \alpha\Delta T \,. \qquad (4.1.1)$$

At high enough frequencies, the temperature within the element does not have time to come to equilibrium with the current, and so a temperature difference does not develop across the device. In this case

$$V_{fast} = V_{Ohmic} = IR \,. \qquad (4.1.2)$$

The basic idea is that at different time scales, the Ohmic and Seebeck contributions to the measured voltage can be separated.

Let us consider the dc situation more carefully. The heat balance at the top metal contact to the device can be written according to Eq. 2.2.1, where P_{top} is the heat exiting the top of the element from the environment.

$$\begin{aligned} -P_{top} &= P_{cond} + P_{Pelt} - P_{Ohmic} \\ -P_{top} &= \frac{\kappa\Delta TA}{L} + \alpha T_{top} I - P_{Ohmic} \end{aligned} \qquad (4.1.3)$$

Likewise, the heat exiting at the bottom of the device can be written:

$$P_{bottom} = \frac{\kappa \Delta T A}{L} + \alpha T_{bottom} I + \frac{1}{2} I^2 R_{int} . \quad (4.1.4)$$

If the two previous equations are summed, the Ohmic heating term cancels out. After doing this and scaling the equation with the factor $\frac{\alpha\sigma}{2I\kappa}$, we obtain

$$\frac{\alpha^2 \sigma}{\kappa} \frac{T_{top} + T_{bottom}}{2} + \frac{A\alpha\sigma}{LI}(\Delta T) = \left(P_{bottom} - P_{top}\right)\frac{\alpha\sigma}{2I\kappa} . \quad (4.1.5)$$

In this expression, we can recognize that the left-most term is simply ZT, the thing we want to measure. Assuming that $R = \frac{L}{\sigma A} + R_{contact}$, we can solve for the material ZT in the expression above. We find

$$ZT = \left(\frac{1}{R - R_{contact}}\right)\frac{|V_{Seebeck}|}{|I|}\left[1 + \left(P_{top} - P_{bottom}\right)\frac{L}{2A\kappa\Delta T}\right] . \quad (4.1.6)$$

In the limit where the contact resistance is much smaller than the device resistance ($R_{contact}$=0) and where the sides of the device are adiabatic (P_{top} and P_{bottom} are both zero), one gets that $ZT = \frac{|V_{Seebeck}|}{|V_{Ohmic}|} = \frac{|V_{slow}| - |V_{fast}|}{|V_{fast}|}$, and so one has successfully found the ZT of the material in terms of two easily measured voltage amplitudes. We note that in assuming adiabatic contacts (setting P_{top} and P_{bottom} to zero in Eq. 4.1.6), one is neglecting the conductive heat transport in and out of the electrical contacts to the device as well as any other heat transport mechanisms (convection, radiation). However, symmetric non-adiabatic heat leaks and contact resistances will always decrease the measured ZT relative to the "true" material ZT, and non-symmetric boundaries can be handled (and detected) by performing the measurement with both polarities of current and averaging the results. Therefore, to the extent that the rest of the assumptions needed for the derivation of 4.1.6 hold (e.g. 1-D heat transport), the Harman measurement has the virtue of being conservative; it's dominant error is expected to lead to a less exceptional (smaller) measured ZT.

This turns out to be an accurate way to measure the ZT of bulk thermoelectric samples, provided that the measurement can be made adiabatic or (more commonly) the error due to non-adiabaticity can be explicitly removed [76]. Although the discussion above suggests a measurement in the frequency domain (e.g. sinusoidal current excitation), the measurement is more typically performed in the time domain using a mechanically commutated square current

pulse and a sampling oscilloscope to measure the resulting voltage transient. Nearly adiabatic contacts are attained by using thin wires to make electrical contact, a vacuum chamber to eliminate convective losses, and radiation shields to reduce radiation loss. In macroscopic samples, the error due to electrical contact resistance can be made negligible. By adding thermocouples to the ends of the sample and monitoring the temperature across the device (hereafter called the augmented Harman technique), each of α, κ, and σ can be determined separately.

For small samples, the situation is much more difficult. First of all, the speed at which a thin film sample thermally relaxes is very fast, on the order of L^2/D, where D is the thermal diffusivity. For a 5 μm thin-film whose diffusivity is that of bulk Bi_2Te_3 [77], the time constant is on the order of 20μs. This requires that the measurement of the Ohmic voltage be done at very high speeds, where parasitic electrical effects like wire inductance are also importance, particularly due to the constraints imposed on the electrical wire contacts from the requirement for thermal adiabaticity. In practice, practitioners of the technique use curve-fitting to extrapolate past parts of the data deemed to be invalid due to these effects[78]. These fits are also used to try to remove the influence of the substrate on which the thin-film is grown, which is often impossible or impractical to mechanically remove. The second major problem with the measurement is evident from Eq. 4.1.6. For thin-film samples it is widely appreciated that parasitic electrical contact resistances typically constitute a significant fraction of the total resistance of the sample. If great care is not taken to measure the contact resistance and remove it from the measurement, this will result in an incorrect estimate of the ZT of the material.

To investigate the severity of some of these issues, we developed a customized finite-difference time-domain approach to solve the Onsager equations for thermoelectric transport in the time domain using a Crank-Nicolson discretization scheme accurate to second order in the time and space discretization steps. The physics we simulated is captured by the two equations:

$$C\frac{\partial T}{\partial t} = \nabla\cdot\left(\kappa\nabla T\right) + \frac{J^2}{\sigma} - JT\nabla\alpha$$
$$E = \frac{J}{\sigma} + \alpha\nabla T \qquad (4.1.7)$$

together with the relevant boundary conditions on heat and current flow. **Figure 51** shows the results of a simulation of the transient Harman measurement on a 10 micron thick

thermoelectric thin film atop a 90 micron substrate. A number of difficulties arise compared with the measurement of a single thick bulk thermoelectric material.

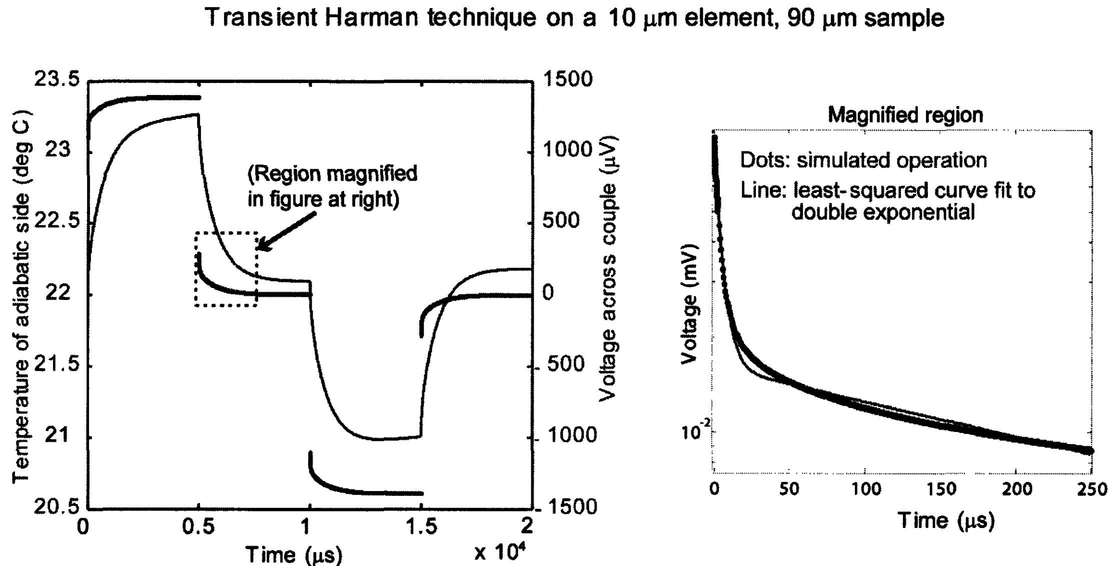


Figure 51. The thermoelectric response of a 10 micron thin film on a 90 micron substrate during a simulated transient Harman measurement. The figure on the right is a zooms in on the region of time immediately after the current pulse is turned off.

First, because the film is very thin the aspect ratio of the test sample is usually quite flat compared to the long aspect ratios that are used for traditional bulk Harman measurements. As a result, the temperature difference for an applied current is smaller, making the resulting thermoelectric voltage smaller as well. The fact that the thin-film in this example (e.g. nanodot ErAs) was on a substrate (e.g. highly doped InP) with its own thermoelectric properties further complicates the measurement, and curve fitting must probably be used to separate the effects of the two materials. As mentioned earlier, the response time of the thin-film voltage is in the microsecond regime due to the speed with which heat can exit the film. While a high-speed measurement introduces another level of complexity to the measurement, one might hope that this eases some of the difficulty in separating the response of the thick and thin film due to the different time scales with which the (slow) substrate and (fast) thin-film thermal relax once the probe current is shut off. This is not necessarily so. As shown in the rightmost part of **Figure 51**, the thermal relaxation is a multi-exponential process because of the distributed heat capacity of the thin-film and substrate so that a simple single (or double) time-constant exponential model does not describe the time dependence of the thermal behavior.

An additional serious issue that we feel has not received much attention (and was not addressed in **Figure 51**) is the effect of non-adiabaticity on the measurement. A traditional transient Harman measurement is conducted with one or both sides of the generator in adiabatic (or near adiabatic) conditions like those we assumed in the theory above. A great deal of effort of has gone into modeling the necessary corrections to the measurement due to the effects of non-adiabatic heat loss even for macroscopic TE elements suspended by wire contacts in vacuum[68]. Thin-film transient Harman measurements have one side of the element on a heat sink (e.g. a substrate) and a top side intended to be adiabatic. Unfortunately, the electrical lead required on the top side of the element (together with the requirements that the lead be low resistance and low inductance) means that the thermal contact to the top of the element is not really adiabatic. In general, both the "heat sunk" and the "adiabatic" side of the film are coupled to the thermal environment with frequency dependent thermal resistances that depend on the detailed nature of the contacts. The contrasting thermal parasitic resistances for a traditional macro-scale transient Harman measurement and a thin-film measurement are shown below in **Figure 52**.

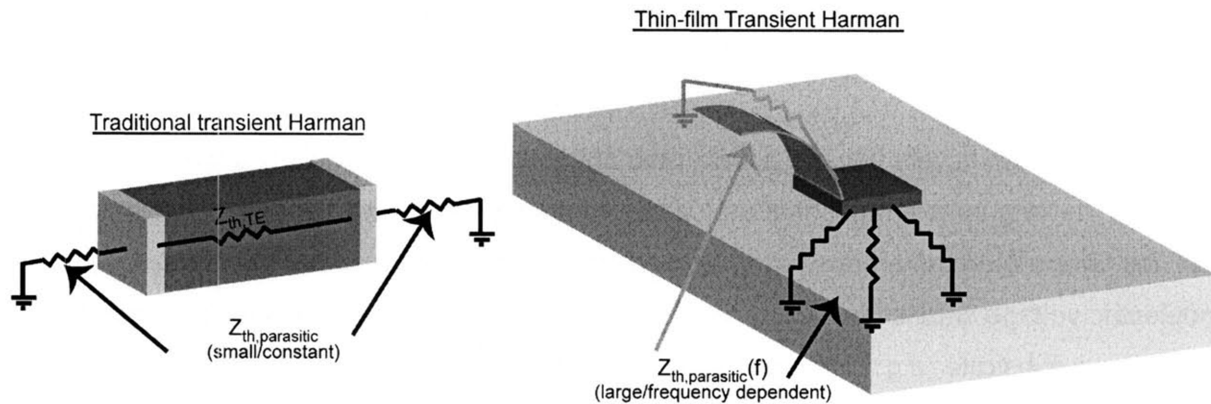


Figure 52. Important parasitic resistances for conventional transient Harman measurement and thin-film transient Harman measurement.

To get a quantitative sense for the importance of the heat transfer through the contacts, we can compare the size of the parasitic thermal resistance of the top lead and the parasitic thermal resistance of the substrate with the thermal resistance of the film itself. We choose a representative TE element that is 5 μm thick, 150 μm x 150 μm in area, and has a thermal conductivity of 0.005 W/cm^2/K. Using thermal quadrupoles, we model the substrate as a semi-infinite half space made of silicon, and we model the electrical contact as a gold ribbon 5 μm in

thickness and 150 μm in width (thermally, a semi-infinite fin) [55]. The resulting thermal resistances are shown below in Figure 53.

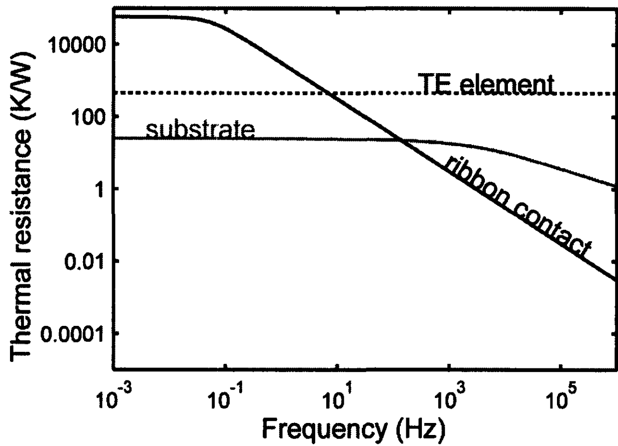


Figure 53. A plot of the two contact thermal resistances present during a thin-film transient Harman measurement, as well as the thermal resistance of the TE element.

As desired, the substrate thermal resistance is a good (but not perfect) heat sink whose thermal resistance is not more than 5% of that of the TE element. The thermal resistance of the gold ribbon electrical contact is very large for low frequencies; since we have assumed it is semi-infinite, the only way heat can leave the ribbon is through natural convection into the environment (set by assuming a heat transfer coefficient of 0.0002 $W/cm^2/K$). Thus the contact is effectively adiabatic at low frequencies, as desired. However, at time scales below around a second (for the geometry chosen here), the ribbon's thermal resistance drops below that of the element, and its effects on the heat transport and on the temperature drop across the element can no longer be ignored. Unfortunately, a transient Harman measurement on a thin-film is necessarily performed in this high frequency regime. For example, in the work of Venkatasubramanian et al.[22], the voltage response of the thermoelectric element was probed down to the 1 μs time scale in order to extract their value of ZT=2.38, and (due to a parasitic inductive spike from their electrical contact), they report that it was necessary to extrapolate to get the correct voltage for times under 1 μs. The magnitude of the non-adiabatic problem depends on the details of the film and contact geometry and materials, but it is not clear to us how these considerations are to be accounted for when thin-films are measured using the transient Harman method, particularly if extrapolations to different time-scales are performed as part of the transient measurement, or if curve fits are attempted using different TE element mesa sizes.

Finally, non-uniform current injection across the device area may be a problem with the transient Harman technique applied to flat mesa structures, although we have not done (or seen) very quantitative modeling of the effect. At the very least, the assumption of an equipotential along the top contact must be carefully checked for the measurement to be valid. For example, in **Figure 52**, the current injected into the ribbon may not spread evenly across the top contact, but may instead bunch up near the ribbon. This again depends on the specific geometry and quality of the electrical contacts, but it is a serious challenge to design a good electrical contact (one that enforces an equipotentials on the ends of the TE element) that does not also act as a heat sink, particularly at high the thermal frequencies needed to perform the transient Harman measurement on thin-films.

Oddly, these issues do not appear to have been discussed much in the literature, despite the fact that this technique was recently used to obtain the most celebrated (and record-breaking) thin-film ZT figures to date[22]. We mention them here to try to make the point that the thin-film Harman measurement is a qualitatively different measurement than its bulk counterpart, and therefore should be held up to the same exacting standards as any new measurement.

DC measurement techniques

A steady-state measurement of a thermoelectric typically applies a fixed temperature difference across the device and monitors the resulting electrical and thermal transport [79]. Determining thermoelectric characteristics this way requires more independent voltage and temperature measurements than the Harman technique, but it has a couple of advantages. By measuring a device in a dc fashion closely resembling a power generation (or cooling) application, one might reasonably expect that the result will be a better representation of the performance of the device. Also, sources of error in the transient Harman technique (e.g. non-adiabaticity due to thermal radiation) can become more severe at high temperature, and the dc technique does not require adiabatic boundary conditions on the element.

A straightforward method for steady-state TE characterization is to place the open-circuited device between hot and cold thermal sinks and simultaneously measure the voltage across the sample, the temperature across the sample, and the heat flux through the sample. Assuming uniform sample properties and a small temperature difference, an open circuit voltage measurement from contacts defined at the end of the sample gives the Seebeck coefficient

directly. The electrical conductivity of the sample can be obtained from a high speed resistance measurement, or by closing the circuit by connecting the element to a variable resistor. In this case, the generated power from a small, stable temperature difference across the element is measured and the generated power is maximized by varying the load[80]. At maximum power, the load matching condition holds, and the resistance of the load will match the total series resistance of the device.

One can determine the thermal conductivity directly from the basic relation $Q = \frac{\kappa \Delta T}{L}$. An excellent discussion of this technique for measuring thermal conductivity has been given by Tzeng et al[81]. One of the measurements discussed there is depicted below in Figure 54. This method has also been successfully incorporated into a general thermoelectric test setup[80]. We will adopt a variant of this method for the measurements later in this chapter.

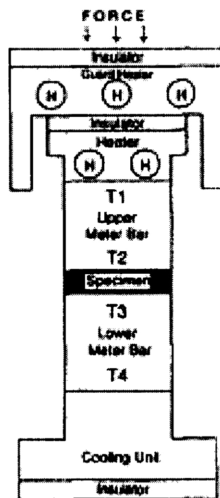


Figure 54. Thermal conductivity test-fixture (Tzeng, 2000)

In this measurement, the sample is squeezed between a heater and a cooling plate using a known pressure (important for insuring repeatable thermal interface conductance). Two meter bars of known geometry and thermal conductivity are positioned immediately above and below the sample. Thermal transducers (thermocouples, resistive thermometers, or thermistors) are positioned inside the meter bars in at least two positions. If the area of the measure bar is well-

matched to that of the sample so that the heat flow through the sample is truly one dimensional, then both the heat flow through the sample and the temperature at the very edges of the sample can be accurately determined (through linear extrapolation of the temperature profile to the end of the rod), allowing the thermal conductivity to be determined precisely. By operating the meter bars for a small temperature difference centered on a temperature T, the full temperature dependence of the sample's thermal conductivity $\kappa(T)$ can be determined.

This measurement is well suited for larger samples, but (as we shall see) for thin/small samples several practical difficulties arise. The typical dimensions of a thin-film thermoelement are around 2-100 μm of thickness and about 1 mm^2 of area. Most thin-films are grown on a substrate which provides mechanical stability, and once again the effects of this substrate will be lumped in with the measurement. If the either the dc or the transient methods above are to be employed for thin-film characterization, some means of separating the effects of the substrate and the film must be provided. The method shown above calls for the construction of a meter bar of the same diameter of the device to ensure simple linear transport, but for small samples that may be impractical. One may use a meter bar which is slightly thicker than the sample, and if the positions of the thermocouples T2 and T3 are positioned far enough from the sample for the heat flow to assume a simple linear profile in the meter bar, the bars can still be used to measure the heat flux accurately. Unfortunately, in this geometry knowledge of the temperature at the interface has been sacrificed, since linear interpolation of the meter bar temperature profiles will no longer yield the correct temperature. One is forced to obtain the temperature across the device using a thermal transducer placed as close to the interface as possible.

There are other problems with the meter bar method for thin-films. Parasitic thermal convection/radiation out of the bar is a source of error in the measurement, and becomes worse as the surface-area to volume ratio of the meter bar increases, imposing a limit on the minimum thickness of the meter bar. The non-negligible size of the thermal transducers in the bar (and the holes in the bar where they rest) also create practical issues. Commercially available thermocouples are presently available with wire diameters greater than 25 μm. Smaller thermocouples can be made, but the issue of ensuring that they are in good thermal contact with the sample then becomes more difficult. The parasitic heat conduction into the transducer wires and the distortion of simple linear heat transport within the meter bar are both worse for smaller bars.

Specialized techniques for thin films

To circumvent the difficulties encountered when trying to apply either of the standard macroscopic TE characterization methods discussed, methods have been developed specifically for thin-films.

For thermal conductivity measurements, the so-called 3ω technique is the measurement of choice for thin-films[57,82]. On an insulating thin-film, the technique is performed by patterning a resistive line (typically Pt, or another metal with a well-characterized temperature coefficient) on top of the film to be measured. A pure sinusoidal current is driven through the metal line at a frequency of ω, resistively heating the sample and producing a temperature wave downward into the film and substrate at a frequency of 2ω. That temperature variation produces a resistance variation in the metal line according to the resistive temperature coefficient of the metal. This modulated resistance mixes with the applied current signal and gives rise to a small component of the voltage across the resistor at a frequency of 3ω. The frequency spectrum of this 3ω voltage, combined with knowledge of the substrate thermal conductivity and thin-film thickness, can be used to extract the thermal conductivity of the thin-film. A circuit to do this is shown in Figure 55, and an example of the output of a measurement is shown in Figure 56 (to further extract a thermal conductivity from the measured temperature, one must use a specific resistor geometry). For thermoelectric samples, the thin-film must be topped with a thin thermally conductive electrical insulator like SiN, so that the heater current is confined to the metal[83]. In practice, this is not a serious problem for the measurement because the thermal conductivity and diffusivity of the thermoelectric thin-film is so much lower than the insulating layer. If the substrate thermal conductivity is also unknown, then by measuring the thermal conductivities of several superlattices of varying length, an attempt may be made at extracting this as well. If the width of the heating wire is varied, the cross-plane and the in-plane thermal conductivity can be separated[84,85] with a careful fit.

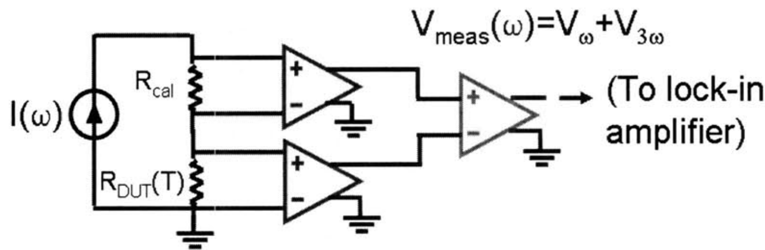


Figure 55. Circuit used to perform 3ω measurement. To measure $V_{3\omega}$, the voltage across the device (DUT) is subtracted from a linear reference voltage to increase the avalable dynamic range of the lock-in.

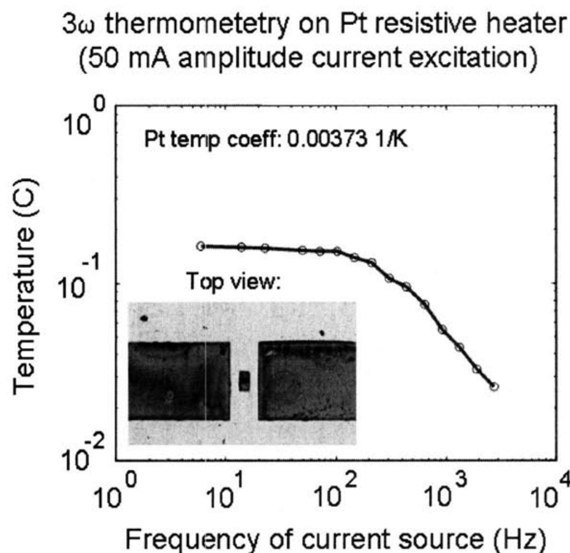


Figure 56. Example of 3ω measurement, performed on a platinum resistive heater suspended on a silicon membrane (see inset). The frequency dependence of the temperature resembles a single-pole transfer function with a pole corresponding to the thermal capacitance of the silicon membrane.

The TLM ("transmission line measurement" or "transfer length measurement")[86] method determines in-plane conductivity of a thin-film with measurements of the resistance between varyingly spaced contacts. If the sheet resistance is known (or measured by standard test structures), the contact resistance can be extracted from this technique. This idea was extended to the cross plane direction by using varying length superlattice samples by Yang et al.[87]. However, separating the contact resistance from the "true" material resistance was problematic except at high temperatures where the material resistance was large relative to the contact resistance. Specialized structures have been designed for measuring highly conductive

thin-films in the cross-plane direction[88], but have not been applied to thermoelectric thin-films, as far as we know. Measuring the cross-plane electrical resistance in a TE thin-film remains a very difficult experimental challenge due to the high conductivity and importance of the contacts. Even in-plane measurements of the resistivity can yield errors due to the parasitic conductivity or Seebeck coefficient of the underlying substrate, especially at higher temperatures when the substrate might inadvertently become intrinsically conductive. A more insidious error can arise due to the Seebeck effect in the thin-film itself, and is shown in Figure 57.

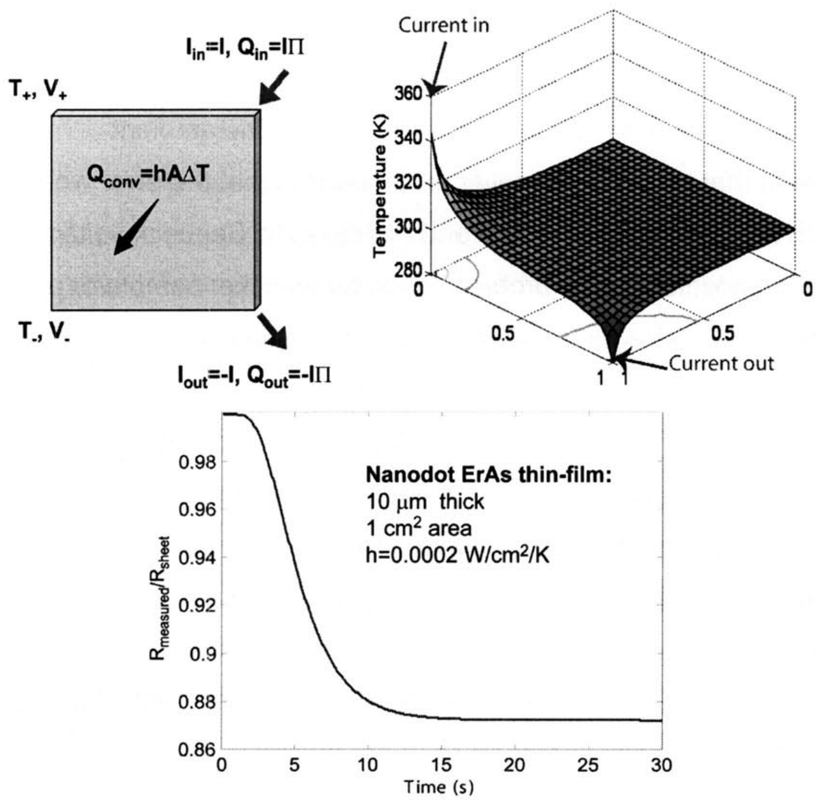


Figure 57. 2-D time domain finite difference simulation of a Van der Pauw in-plane electrical conductivity measurement on a suspended thin-film. The probe current can significantly perturb the sample temperature, adding a spurious thermal voltage to the measured Ohmic voltage drop.

The sheet resistance of a thin-film Van der Pauw test structure (like the square sample depicted in the upper-left corner of Figure 57) is typically measured by injecting and removing current from one side of the device while monitoring the resulting voltage across the other side. We simulated the time-dependent 2-dimensional temperature and voltage profile by solving the

Onsager relations over a 2-D square thin-film subject to a constant heat transfer coefficient off of the surface of the film:

$$C\frac{\partial T}{\partial t} = \nabla\bullet(\kappa\nabla T) + \frac{J^2}{\sigma} - JT\nabla\alpha - \frac{h}{\tau}(T - T_{amb})$$
$$E = \frac{J}{\sigma} + \alpha\nabla T \qquad (4.1.8)$$

We found that problems with the measurement can arise if the injected current perturbs the temperature of the sample (as shown in the temperature plot in the upper right-hand corner of Figure 57) so that a spurious Seebeck voltage is added to the Ohmic voltage measurement. The plot at the bottom of Figure 57 depicts the error in the measured sheet resistance that develops after a step probe current is applied. The situation actually becomes worse in a more adiabatic environment so that a measurement in vacuum will show a larger error than the ambient conditions shown here. It might seem that by using a smaller probe current that the error would disappear, but this is incorrect; both the Ohmic voltage drop and the parasitic Seebeck voltage scale the same way with current. One solution to this problem is to use a thicker sample or a substrate; both act to "short out" the undesirable temperature difference, mitigating the resulting Seebeck error.

There have been a few specialized techniques developed for measuring the Seebeck coefficient of thin films. As was the case with thermal conductivity measurement, the chief problem here is accurate determination of the temperature across a small distance. For a valid Seebeck coefficient measurement, one must also ensure that the isothermal planes of the measured temperature difference correspond to well-defined voltage equipotentials. In-plane thin-film Seebeck measurements can be made using similar geometries to the macroscopic dc measurements described above, since long (~mm) samples can be prepared[85]. There is one possible problem that we illustrate using the setup pictured below in Figure 58. Here two thermocouples simultaneously monitor voltage and temperature.

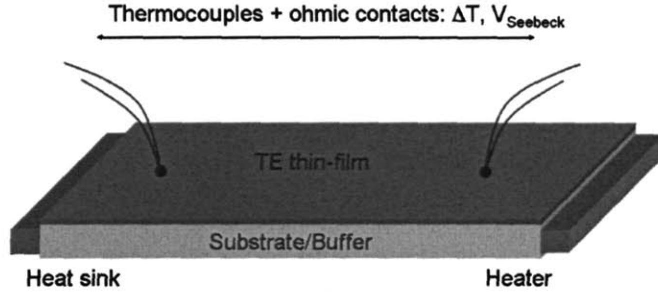


Figure 58. Sample geometry for in-plane Seebeck coefficient determination.

For the measurement to work, the heating and the cooling at each end of the sample strip must be allowed to determine flat isotherms, such that thermal transport down the strip in the region between the monitoring thermocouples is effectively one dimensional. If the substrate or buffer layer beneath the thin-film of interest is electrically conductive, an additional complication arises. Because the Seebeck coefficient of the buffer is usually different than that of the TE thin-film, the same temperature difference across both will give rise to different voltages, leading to circulatory electrical currents between the two layers to cancel out that potential difference. This would give rise to a spurious Ohmic contribution to the measured Seebeck voltage which could either add or subtract from the real Seebeck voltage. So in this situation, care must be taken to properly account for the effects of the buffer on the measurement using modeling and separate measurements on the properties of the substrate/buffer.

Cross-plane measurements of the Seebeck coefficient are complicated by the same issues. One version of the measurement relies on a metal resistive heater on top of a thin-film TE mesa (or otherwise well defined area for cross plane transport) which can simultaneously monitor the temperature of the top contact of the superlattice through its temperature coefficient. To remove the contribution of the substrate/buffer Seebeck coefficient from the measurement, there are two methods. One can model the three dimensional heat spreading and voltage through the substrate, neglect any resulting thermoelectrically induced internal currents, and then fit parameters to that model using a circuit theory analogy for the thermal resistance network[89]. This is technically valid only if the isotherms and equipotentials are aligned in the test structure, and is preferably done with several different length thin-films to aid the fit. Or one can design the sample geometry so that the measured equipotentials and isotherms are as close as possible to one-dimensional heat transport, and pattern two samples for measurement, identical except that one is absent the thin-film[83,87]. Using a subtractive technique, one can then

remove the Seebeck contribution of the buffer/substrate. An example of this is shown in (Figure 59) [90]. Unfortunately, both techniques are somewhat mathematically and physically complicated, involve patterning special test structures, and require the accurate knowledge of the thermal conductivities of the involved materials.

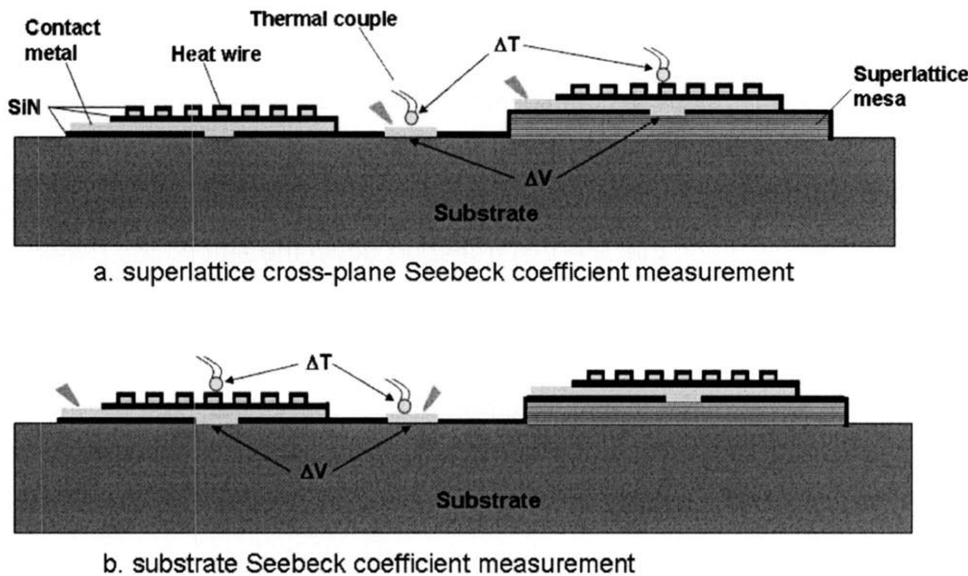


Figure 59. A schematic of a cross-plane Seebeck coefficient measurement (Zeng et al., 2005)

In summary, there are a few key difficulties in the thermoelectric measurements on thin-films. First, accurately measuring the temperature at a precise location on scales of less than 1mm is difficult and often constrained by the properties of the contacts and the temperature transducer. Second, for practical reasons it may not be possible to locate the thermocouple exactly at the interface of interest, adding an unknown interface thermal parasitic to the measured results. Third, test structures may need to be grown whose geometry may not give rise to the simple one-dimensional heat and current flow desired for straightforward interpretation of the measurement results. Finally, in all measurements a way must be found to separate the Ohmic and Peltier contributions when making voltage measurements on the sample. In the remainder of this chapter we will explore the limits of DC measurements to overcoming these challenges. In Chapter 5 we will present an approach to microscale thermal measurements that circumvents some of these issues.

4.3 High temperature power generation test benches

The goal of this chapter is to present instruments designed to address the key problems of thermoelectric element and generator characterization and to explore the range of measurements that can be performed successfully with the instruments. The motivation for designing the instruments is straightforward: in order to test and compare thermoelectric generators it is necessary to have a repeatable, flexible, and accurate metrology platform capable of testing the system under conditions approximating those of real operation and recording all of the information pertaining to the generator's operation. Also, before a generator can be built it is necessary to test and compare the performance of different thermoelectric materials to explore their suitability for power generation. The platform that was developed out of this thesis work is capable of both of these tasks.

Two generations of test equipment were developed as part of this thesis. The first was a high-temperature difference platform designed to quickly test the power generation capability of single elements made from thin-film materials and of packaged generators. The second was a vacuum system capable of higher temperatures, more precise and repeatable measurement conditions, and the added capabilities of measuring heat transfer and power generation efficiency. In this section we discuss each system in turn.

4.3.1 Rapid thermal test platform

The system is depicted below in Figure 60. The sample sits between two 1 cm^3 OFHC (oxygen free high conductivity) copper blocks. The copper has excellent thermal conductivity (κ=3.91 W/cm/K) and very high electrical conductivity. Two thick wires (18 AWG) inject and remove the generated current from the top and bottom of the sample, respectively. The temperatures at the edges of the copper blocks are monitored using two type K thermocouples soldered into holes within 1 mm of the edge of the block. The bottom of the sample is connected through the copper to an extruded aluminum cold plate (LC high-flow plate, by Solid State Systems), whose heat transfer coefficient was measured in Section 3.5 (h=0.66 $W/cm^2/K$). Heat is injected using a high-temperature cartridge heater (Chromalox Incoloy-sheathed, Omega Engineering) and an analog voltage controllable solid-state relay. The temperature is set and stabilized using a third thermocouple adjacent to the heater with a Labview PID controller written for that purpose. The entire system is mechanically aligned using ceramic rails.

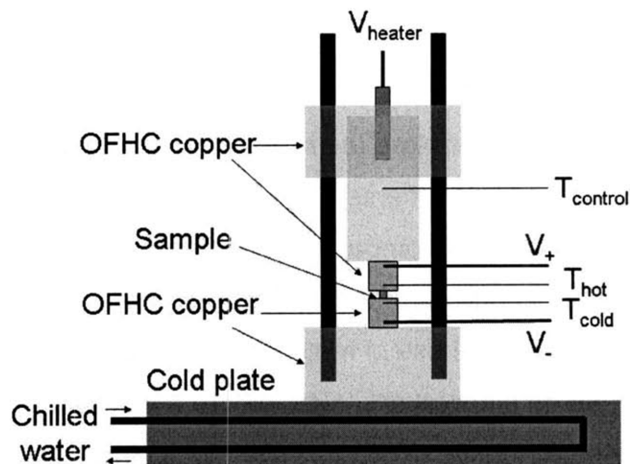


Figure 60. Experimental setup for measuring generated power density.

The system was originally meant to operate at temperatures as high as 830 degrees C, limited by the maximum sheath temperature of the cartridge heater. However, due to parasitic convection and radiation into the environment the maximum practical temperature is around 450 C. This is partly related to the fact that when exposed to air at temperatures above around 400 C the copper heater block oxidizes, developing a flaking black coat of copper oxide. This is an annoyance because it greatly increases the emissivity of the copper, further increasing the heat loss to the environment. The tests in this open air system were therefore kept under 400 deg C. The next generation test bench was designed to operate in vacuum, which eliminiated this problem. Another solution that has been used by others in the field is plating the copper (whose high thermal conductivity and machinability are very desirable) with another metal like nickel that is less prone to oxidation damage at high temperatures.

4.3.2 Milli-Ohm matched load for power measurements

Once a large temperature difference can be applied across the device the open circuit voltage (the voltage resulting from the Seebeck and Thomson effects) can be measured. This in itself may be very useful. For small temperature differences, the open circuit voltage divided by the temperature difference is the Seebeck coefficient at the average temperature of the material. However, for elements used in power generation applications the large temperature variation across the element in this measurement will generally cause the material Seebeck coefficient to vary significantly with temperature (and hence position) in the device and therefore the open

circuit voltage is not trivially related to the Seebeck coefficient at one temperature. (Although if one knows the full temperature dependence of the Seebeck coefficient, one can feed that data into a simulation such as the finite difference model discussed in Chapter 2 and compute the open circuit voltage analytically.) But since "large ΔT" open circuit voltage ultimately drives the power generator performance, measuring the open circuit voltage under large temperature differences makes sense from the standpoint of understanding the potential of a thermoelectric element for power generation.

Beyond simply measuring the open circuit voltage with large temperature difference it would be useful to also measure generated power directly. There are a few reasons why this is useful. First, the same property variation with temperature that we just discussed for the Seebeck coefficient also occurs for the electrical conductivity of the material. Because the open circuit voltage is measured under conditions of zero current flow, the measurement can provide no information about the conductivity of the thermoelectric element. To get at that information one must close the circuit and allow current to flow which in practical terms means allowing the element to generate power. Aside from the pure material conductivity (and conductivity variation with large applied temperatures) the parasitic electrical resistances associated with the element contacts are also captured in a power generation measurement. Since these can also display substantial temperature dependence (low contact resistance, high temperature Ohmic contacts are notoriously difficult to make in practice) it is important to define a measurement that also captures their influence. There is no measurement on a single thermoelectric element which is more predictive of its performance in a power generator than a power generation measurement on the element itself.

An additional metrology-related reason also exists for measuring true power in addition to open circuit voltage. When making an open circuit voltage measurement on an element with a large temperature difference there are sometimes parasitic open circuit voltages that can contribute errors. But when the circuit is closed these parasitics can be immediately recognized since they will typically degrade the power generation of the element (even though they may have increased the open circuit voltage). For example, copper oxide is known to have a high electrical resistivity, a relatively poor thermal conductivity, and a large (> 500 μV/K) Seebeck coefficient. If a thin layer of copper oxide is inadvertently included in a measurement of the open circuit voltage across a thermoelectric element (for example, at one of the element-to-copper block interfaces) the low thermal conductivity of the oxide may result in a substantial

temperature drop across the oxide interface. The problem with this is that owing to its high Seebeck coefficient the oxide can then contribute a significant spurious voltage to the open circuit voltage measurement which could go undetected if the experimenter is not careful. If the same element is tested for power generation it will become immediately obvious that something is wrong since the high electrical resistivity of the copper oxide will decrease the generated power (and change the expected impedance matching condition) for the element.

The most serious impediment to generating power from a single element is impedance matching. Thermoelectrics are often semi-metals or heavily doped semiconductors, and typically possess electrical conductivities in the neighborhood of 500 1/Ω/cm. It is often inconvenient for the element to be much smaller than 1 mm^2 in area due to issues with cleaving, handling, and contacting the device. The thickness of the devices (including any substrate used to stabilize thin film materials) can be on the order of 200 μm or less. This means that typical single TE element resistances are in the milliohm regime. The optimum load will be equal or on the same order, as discussed in Section 2.7. The generated power drops off roughly linearly as the load resistance becomes much larger than the internal resistance of the element resistance (Eq. 2.5.1 and Eq. 2.5.3). For a load $R_L >> R_{int}$, the generated power is reduced by a factor of R_{int}/R_L. It is desirable to reduce this mismatch factor as much as possible and closely approach the optimum power generation bias point.

Previous efforts in this area relied upon switched banks of precision milliohm resistors to synthesize small resistive loads[91]. To obtain continuously varying small impedances, we made a variable load using the linearized transconductance of a power transistor.

The load used across the sample is shown above in Figure 61. The first part of the load is a precision four-wire current monitoring resistor from Ohmite, typically 5 mΩ in value. The voltage across the resistor is read using an HP digital multimeter to obtain the current through the sample, and the voltage across the sample can be directly monitored with another voltage measurement. Two STV160NF02L power FET transistors are used in parallel to synthesize an arbitrary additive load. The drain-source on resistance of a single FET is under 2.5 mΩ according to the specifications of the device; here two were used in parallel to lower further the minimum conductance. By changing the gate-source voltage the load seen by the thermoelectric sample can be tuned from just over 5 mΩ on up to an effectively open circuit. In addition to the resistances shown in Figure 61 the sample also sees several mΩ of parasitic

wire and contact resistance in series with the load. This sets the upper limit on impedance matching with a single element. The temperature of the FETs was stabilized using a thermoelectric stage to help alleviate any power dissipation dependent nonlinearities. Figure 62 shows a plot of the V_{DS} vs. I_{DS} for various gate voltages. It is apparent that the linearity of the curves at low V_{DS} is quite good, although at higher V_{DS} (not shown here) the curves begin to roll over as the FET approaches saturation.

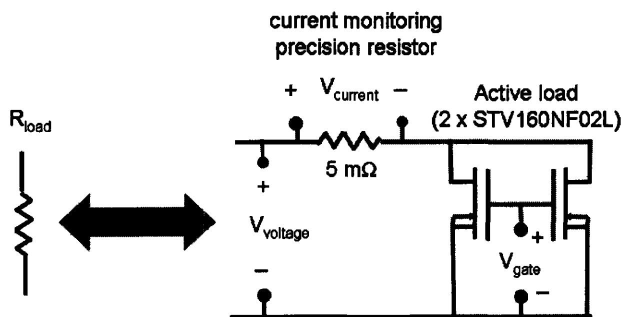


Figure 61. Active load used for the measurement of the device current and voltage. In preliminary measurements, an initial resistance calibration was used to dispense with the precision current sensing resistor.

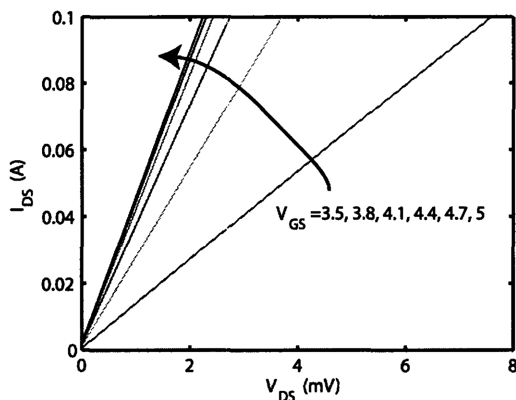


Figure 62. I-V curves of active load, taken with semiconductor parameter analyzer. A parasitic series lead resistance of around 20 mOhms is also included in the circuit.

The equation describing the operation of the active load is just the equation for the current through a FET (since we can treat the two identical parallel FETs as a single FET with twice the transconductance):

$$I_D = 2kV_{DS}\left[(V_{GS} - V_T) - V_{GS}/2\right]. \qquad (4.3.1.1)$$

It is clear from this expression that the drain-to-source resistance of the FET (V_{DS}/I_{DS}) is itself a function of both V_{GS} and V_{DS}. The former is desirable, since it allows us to control the resistance of the load using the gate voltage. The latter is an annoyance, and although it should not ideally contribute an error (since we are always measuring both the voltage and the current through the active load, we will always know its equivalent impedance), it causes the resistance seen by a thermoelectric element to change as the temperature applied to the element (since the voltage displayed across the active load, which is V_{DS}, is changing).

The circuit of Figure 61 was used for the power measurements in Section 4.3. Subsequent measurements made use of an improvement shown below in Figure 63.

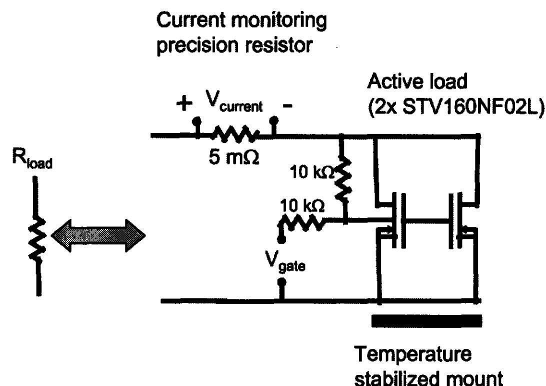


Figure 63. Milli-Ohm active load with linearizing resistors.

The design in Figure 63 adds two resistors between the drain and the gate to feedback the right voltage from the drain to the gate to linearize the response of the active load. This works because the actual voltage at the gate of the transistors will now be halfway between the voltage of the drain and the applied gate voltage, V_{gate}, so that

$$V_{GS} = \frac{V_{DS} + V_{gate}}{2}. \qquad (4.3.1.2)$$

When this expression for the gate-to-source voltage of the FET is substituted into Eq. 4.3.1.1 above, the terms depending on V_{DS}^2 are cancelled out so that the device behaves as a voltage controlled load. A test of the linearity of the load can be made by fixing the gate voltage to select the desired load resistance and then varying the current through the load with a controllable current source (Figure 64). Over two orders of magnitude in current variation and at

load currents into the Ampere range, the resistance is quite stable. This improved active load was adopted for the subsequent power generation and efficiency measurements later in this chapter.

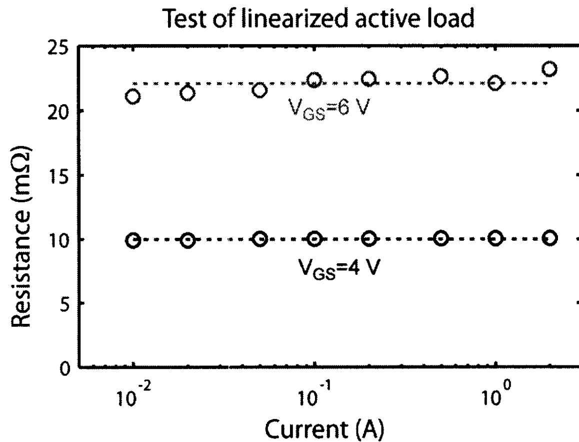


Figure 64. A test of the linearized active load for mOhm load synthesis. For a fixed gate voltage, the current through the active load was varied over several decades with no significant change in load resistance observed.

The active load we have constructed is useful for measuring power in individual thermoelectric elements. It is not meant for storing power or using power for a practical purpose. A packaged thermoelectric generator in the field is typically designed to supply current at an output voltage that as close as possible to the one needed for the application at hand. To stabilize this voltage and perform any needed impedance matching, the first stage seen by the power leaving a packaged thermoelectric generator is typically a DC-DC converter designed to shift the (typically low) voltage level of the generator to exactly what is needed to power a circuit or charge a battery. DC-DC converters capable of handling output voltages as low as 300mV have been designed for thermoelectric applications [92].

4.3.3 Vacuum thermal test station

The power measurement station in described in Section 4.3.1 is useful for quick measurements of generated power and open circuit voltage, but is not capable of measuring generator efficiency because there it has no way of measuring the heat entering (or exiting) the device. Another limitation is that the setup is operated in ambient conditions, so that convection (and the oxidation of the copper heater block) limits the maximum hot side temperature to around 400 C. Another desired improvement to the station was the ability to quantify the contact pressure on

the sample so that the electrical and thermal parasitic contact resistances can be minimized or at least made more repeatable. A new test bench was constructed to address all of these issues, based in part on systems from the thermal metrology literature [80,91]. It is shown below in Figure 65.

High-temperature vacuum test station

Thermoelectric test setup
(inside vacuum chamber)

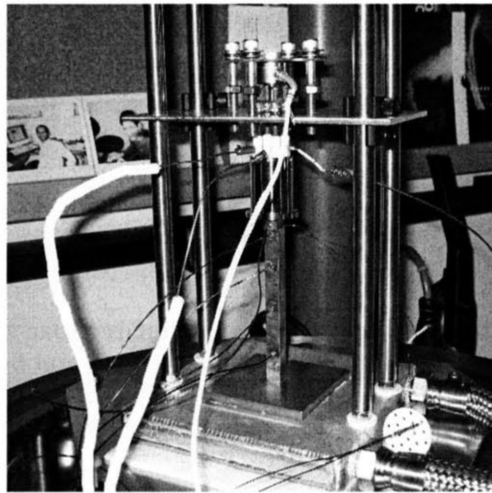


Figure 65. The high-temperature vacuum thermoelectric test bench (a.k.a. Shrek).

The station was built inside of a Kurt J. Lesker vacuum chamber pumped by a Varian V301 turbo-molecular pump backed by a TriScroll 300 oil-free roughing pump. The empty vacuum chamber could reach vacuums in the mid 10^{-7} Torr range in under 30 minutes, and the full chamber could be pumped to around 1 mTorr in under 10 minutes. A high temperature semiconductor wafer heater and temperature controller (HeatWave Labs) were used to establish the hot side temperature for thermoelectric testing. Temperatures of up to 1200 C were possible using the heater. The controller used a thermocouple bonded to the inside of the "button" style heater, and both the heater and the thermocouple were enclosed with a polished molybdenum radiation shield. The cold side of the thermoelectric was set using the water cooled cold plate discussed in Chapter 3, with the appropriate vacuum feedthroughs. It is likely that the imperfect seal on the cold plate is what set the limit of the minimum attainable vacuum pressure with the fully loaded chamber. Nevertheless, the pressures reached (~1 mTorr) were sufficient to eliminate the effects of parasitic heat convection into the air. Electrical feedthroughs were used for the power leads of the heater, for the voltage measurements across the current sensing resistor and across the active load, and for the control signal to the gate of

the active load. The active load itself was positioned inside the vacuum chamber on top of the water cooled heat sink adjacent to the measurement. The proximity of the active load to the measurement was necessary to reduce the parasitic resistance of the copper wires leading from the copper thermoelectric contact blocks to the load.

The voltage measurements from the active load and the current sensing resistor exited the chamber through one of the electrical points, were amplified using Burr-Brown INA106 instrumentation amplifiers, and were read into a computer using a data acquisition board within the LabView programming environment. (The same data acquisition board was used to supply the analog output voltage to tune the active load resistance.) Together the measurements yielded the current and voltage across the active load, and hence the generated power or open circuit voltage.

To obtain the temperature difference across the thermoelectric element or generator two thermocouples were soldered into the center of copper contact blocks between which was positioned the element of generator to be tested. The thermocouples were less than 1 mm from the copper-thermoelectric interface. The lower copper block was actually an extended copper rod that could be used as a meter bar for heat transfer measurements. The basic thermal measurement setup is shown in Figure 66.

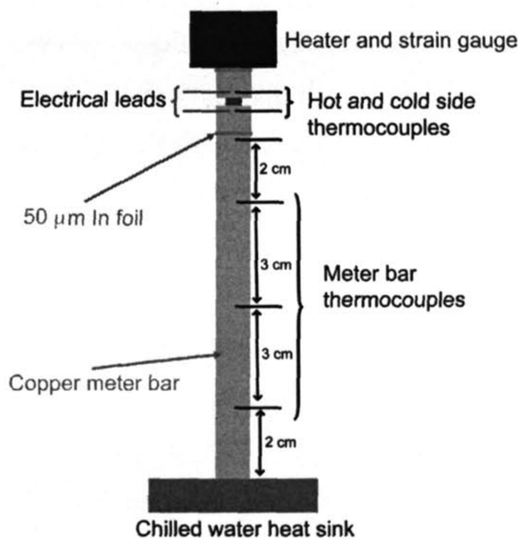


Figure 66. Schematic diagram of the thermal measurements inside of the high-temperature vacuum test bench.

The three meter bar thermocouples were positioned 2 cm from the ends of the meter bar, and spaced from one another by 3 cm. The position of the hot and cold side thermocouples in the copper block was close enough to the element that the temperature drop due to the copper was typically negligible compared with interface thermal resistance (as verified with finite element simulations for specific elements). For example, a 1 mm^2 thermo-element sees less than 1 K/W spreading resistance between the surface of the copper and the thermocouple measuring the surface temperature. As will be discussed later, thermal interface resistances for the same area element were greater than 4 K/W for a liquid metal contact, and greater than 100 K/W for indium or bare copper contacts. All of the thermocouples were K-type so as to be compatible with high temperature operation, and they were passed through a Chromel-Alumel feed-through from the vacuum chamber to a thermocouple monitor (SR560) capable of 0.1 C resolution. The thermocouples were inserted and soldered into holes drilled into the copper bars in order to minimize the thermal interface resistance between the thermocouples and the copper. Two meter bar geometries of identical lengths but different cross sections (circular cross section, area 0.317 cm^2, and square cross section, area 1 cm^2) were built so that elements and generators of various areas and thermal resistances could be accurately measured. Before each measurement, the surfaces of the copper contacts were cleaned and polished.

Once the sample was positioned in the measurement stack, the contact pressure could be set using a piezoelectric strain gauge (Omega). A system of Bellevue (linear force constant) spring loaded set screws was used to set and maintain the pressure. A LabView controller was written to facilitate testing and coordinate all of the simultaneous temperature and voltage measurements. An example of the raw data from a measurement on a thermoelectric generator is shown in Figure 67.

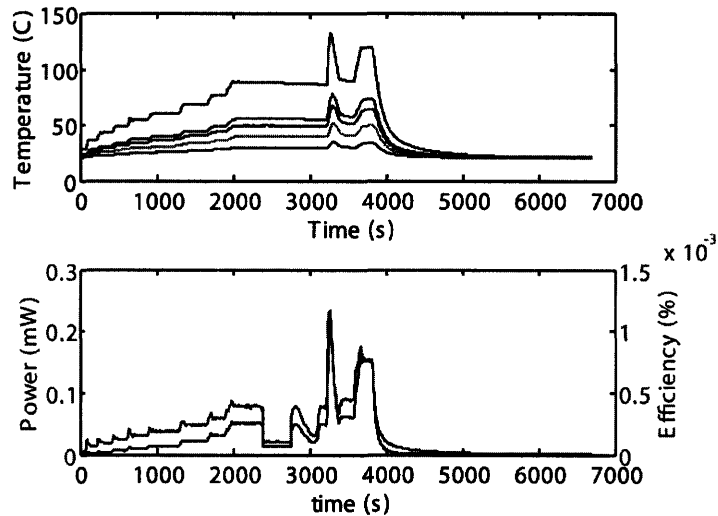


Figure 67. Example of the raw data from a measurement on a thermoelectric generator using the high-temperature vacuum system.

In the uppermost plot of Figure 67, the top two temperature traces are the hot and cold side temperatures of the generator. The lowest three temperature traces are the thermocouple readings from the meter bar. Since the heat transfer in the meter bar is designed to be one-dimensional, and the meter bar thermocouples are equally spaced, the temperature difference between the first and second thermocouple in the meter bar and the second and third thermocouple in the meter bar should be equal. If they are not, it indicates either a source of heat leakage (e.g. convection or radiation from the side of the copper bar) or that the measurement is not in the steady state. To avoid errors due to the latter we found it was necessary to step between discrete temperatures and stabilize the measurement there while data was taken rather than (for example) heating the device up, turning off the heater, and then taking continuous data as it cooled. The time scale required for accurate measurements was generally on the order of minutes or tens of seconds as discussed in detail in Appendix 4.

To verify that the system was working and confirm the repeatability of previous results, open circuit voltage measurements were performed on two samples identical to those that had been earlier measured in the power generation setup described in section 4.3.1. The results of those measurements are in good agreement with one another, as shown in Figure 68.

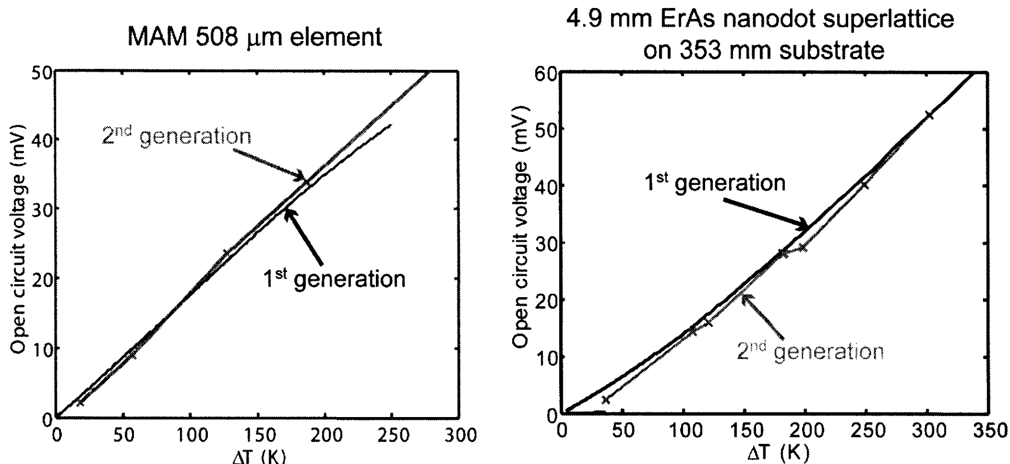


Figure 68. Comparison of the open circuit voltages for two samples measured using the first generation (rapid power measurement) and second generation (vacuum power and efficiency measurement) setups.

A further key piece of information needed to interpret thermoelectric element measurements is the size of the temperature drop that occurs across the interfaces between the copper contact blocks and the thermoelectric element or generator. With the capability of measuring heat transfer afforded by the meter bar in the vacuum test setup comes the ability to characterize these thermal interface resistances. While thermal interface resistances vary somewhat with pressure and the surface morphology of the contacts, the bar chart of Figure 69 serves as a useful estimate for the different interfaces used in this thesis.

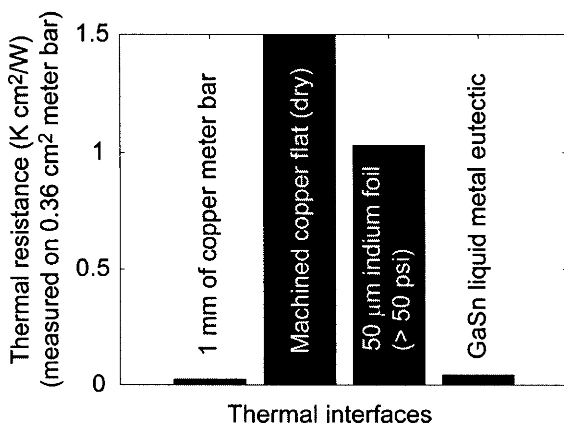


Figure 69. A comparison of different interface thermal resistances used to contact thermoelectric elements.

The baseline thermal interface resistance for the thermoelectric test bench is the thermal resistance provided by the approximately 1 mm of copper between the surface of the copper contact blocks and the thermocouples sensing the hot and cold side temperatures. (The thermal spreading resistance into the copper is not shown here, as it depends on the area of the sample under test). The highest thermal interface resistance was measured between the dry faces of the machined copper contact blocks. At pressures of over 50 psi, the 50 μm indium foil was a slightly better thermal interface. By far the lowest resistance thermal interface was a gallium-tin liquid metal eutectic provided for us by Bob Reeder of Lincoln Labs for this purpose (0.043 K cm^2/W). These values are in line with comparable measurements from the heat transfer literature [93]. We also note that all of the thermal resistances are far larger (> 100x) than the minimum thermal resistance between typical dissimilar solids, given by either diffuse or acoustic mismatch theory (depending on the nature of the interface) [94]. For the copper-copper and copper-indium-copper contacts this is not too surprising since on a microscopic scale it is known that thermal contacts between metal surfaces are strongly affected by the microscopic voids at the interfaces.

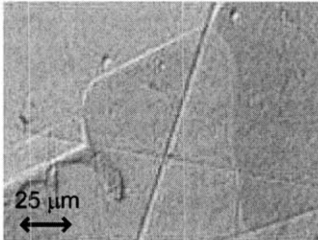
4.4 Bulk thermoelectric elements and generators

4.4.1 Length scaling of Bi_2Te_3-based micro-alloy material (MAM) elements

One of the central themes of this thesis has been the idea that while TE generator efficiency is strongly bounded by the ZT of the material, power can be increased significantly by decreasing the length of the TE elements provided that the parasitic electrical resistance and thermal resistance (including the heat sink) do not intervene. To test this length dependence and to verify our ability to impedance match to milli-Ohm TE elements different thicknesses of thermoelectric material were obtained from Marlow industries. The samples were made of their n-type advanced micro-alloy material (MAM), a multicrystalline material composed of small (1.0μm - 20 μm) grains of material from the Bi_2Te_3 alloy family[95]. A photomicrograph comparison of the MAM material and a traditional thermoelectric crystal are shown in Figure 70. There are two reasons why the micro-alloy material should be more durable than a single crystal. The first concerns the crystal structure of Bi_2Te_3. It can be represented using a hexagonal cell with 5 stacked layers perpendicular to the c-axis (long axis) of the cell. The resulting highly anisotropic structure is known to cleave easily in planes perpendicular to the c-

axis, particularly along the Te^1-Te^1 plane where the bonds are held by some to be van der Waals in character [96]. The MAM material, while still (according to Marlow) preserving some of the beneficial anisotropy in the thermoelectric properties, does not have a single cleavage plane running through as large an area of the crystal as does traditional Bridgeman-grown $B_{i2}Te_3$. This should contribute to the durability. Also, an increased yield-strength of the MAM material (and micro-grained materials in general) is expected from the so-called Hall-Petch relationship, which predicts that multicrystalline materials will have greater yield strength with decreasing grain size (down to the micron-grain size regime). The simplest picture to understand this relation is that the grain boundaries in the microcrystalline material act as barriers to dislocation motion in the crystal, although is some evidence that the story is more complicated [97].

Bridgeman grown, large grain material



Marlow MAM (micro-alloy material)

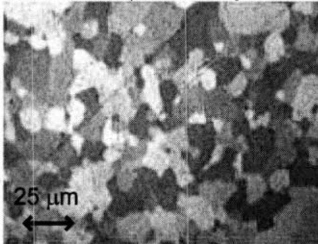


Figure 70. Top photo is a traditional Bridgeman-grown thermoelectric element. Lower photo is the Marlow micro-alloy material (MAM). Photos taken from Marlow website (http://www.marlow.com/AboutMarlow/advanced_materials.htm).

Marlow's manufacturing process for their MAM material is proprietary but we may find examples in the literature of how such materials with similar properties are prepared [98]. According to Marlow, the MAM material is designed to preserve some of the useful thermoelectric anisotropy of the constituent crystals while benefiting from the durability afforded by a smaller grain-size material. This is somewhat surprising since one might expect that the constituent grains making up the material would be randomly oriented, especially if the material is prepared though a powder metallurgical process. However, even if material preparation started with ball-milling a Bridgeman grown Bi_2Te_3 crystal ingot into a powder, the powder particulates might have large

shape anisotropy due to the relative ease of cleavage along the basal planes of the crystal. In other words, rather than a collection of jagged spheroids, one might end up with a collection of platelets. When these platelets are hot-pressed or sintered into a compact, they can stack up so that the resulting microcrystalline material preserves some of the anisotropy of the original crystal.

A calibration of the thermoelectric material properties was provided by Marlow up to a maximum temperature of 337 K = 64 C (since the material was actually optimized for cooling applications). For a large bulk element (0.15 cm x 0.15 cm x 0.19 cm) at that temperature they report α=241 μV/K, σ=657 1/Ω/cm, and κ=0.0153 W/cm/K. The samples were of thickness 177 μm, 254 μm, and 508 μm, and each had identical electrical contacts prepared by Marlow. For the purpose of the study, we made two longer lengths of 1524 μm and 1016 μm by stacking up shorter elements.

During the measurement, the hot side temperature was increased to around 100 C at which point the load across the thermoelectric element was optimized using the active load circuit. A dry contact between the element and the copper block contacts was used with very high pressure (estimated > 100 psi) to force the soft gold contacts to comply with the copper contacts and minimize the electrical and thermal parasitic interface impedance. The active load was used to tune alter the load resistance while the voltage and current from the element were monitored. The resistance giving the maximum power output could then be found and used for the subsequent power generation measurements. Results demonstrating this optimization for thermoelements of varying length are shown in Figure 71. When the optimal resistances are plotted against the length of the elements it is possible to estimate both the bulk conductivity and the parasitic resistance associated with the test leads and the contact resistance (Figure 72) since the optimal resistance for this material is expected to be close to the impedance matched value. A contact resistance of 12 mΩ and a conductivity of 270 1/Ω/cm can be extracted from the curve fit. Unfortunately, the accuracy of the curve fit depends heavily on the reproducibility of the contacts and temperatures for different samples so these values may or may not be indicative of the true material electrical conductivity. For example, a 500 μm thick, 1 mm^2 area sample with the electrical conductivity given by Marlow (657 1/Ω/cm) has a resistance of only 7.7 mΩ, so that if the two contacts to the TE element had a parasitic contact resistance of 3.8 x 10^{-5} Ω cm^2 the parasitic contact resistance would be the same size as the actual

element resistance. While the curve fit we performed is designed to remove the effects of the contact resistance from the resistivity, its magnitude and dependence on contact pressure and interface quality may cause it to vary between measurements.

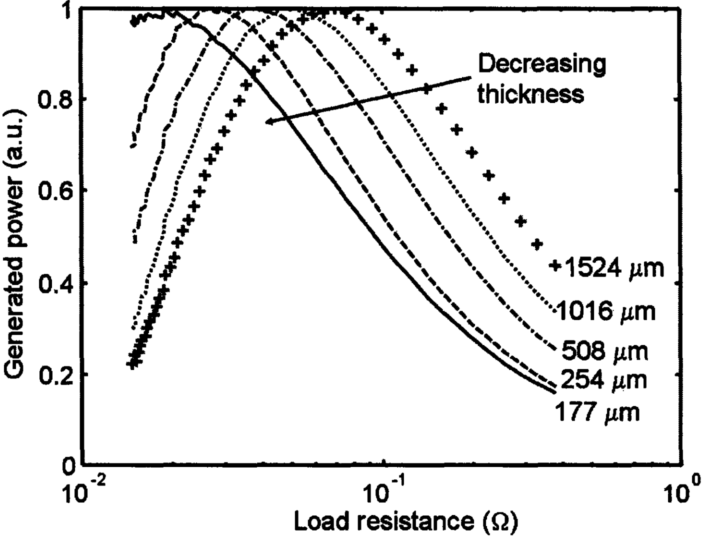


Figure 71. Experimental optimization of the load resistance.

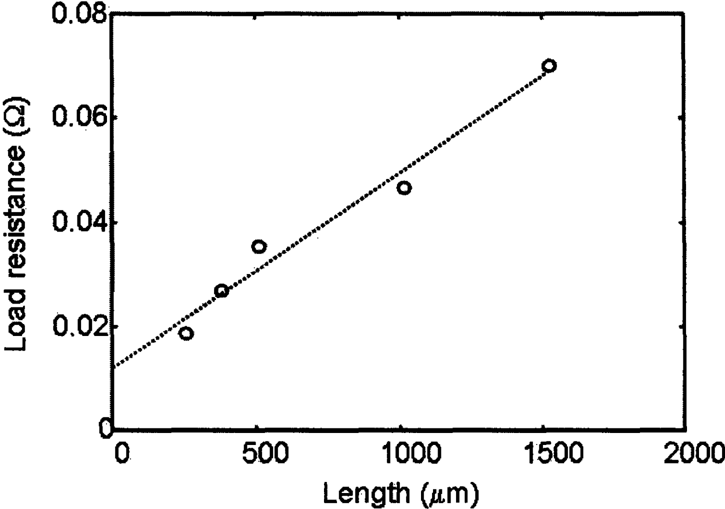


Figure 72. Load resistance vs. length for Bi_2Te_3 elements

Once the optimum load for each element was measured, the heater was used to establish a large temperature gradient across the device. Typically the hot side contact to the element was raised to around 320 deg C, and the cold side contact was maintained below 30 C using the

chilled cold plate. (Higher temperature differences were possible, but they rapidly degraded the material.) By switching the gate voltage the open circuit voltage and the voltage under optimum power generating conditions could be continuously and automatically monitored with a digital multimeter.

An example of the measurements for a 381 μm thick element is shown in Figure 73. The power density is calculated by dividing the power generated by the area of the element (0.01 cm^2). It is clear from the changing slope of the open circuit voltage data that the Seebeck coefficient of the material (which was not optimized for high temperatures) begins to decrease at higher temperatures. This in turn causes the generated power to deviate from its expected parabolic course.

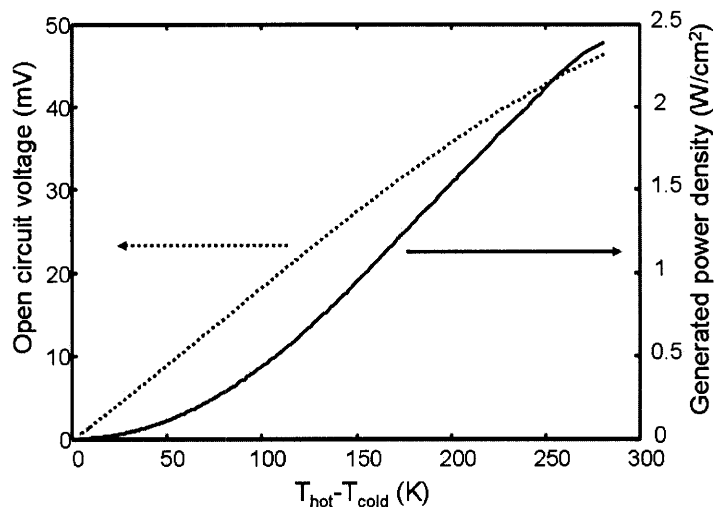


Figure 73. Generated power and open circuit voltage for Bi_2Te_3 element.

Because the electrical resistance of the element is actually a (typically increasing) function of the element temperature the optimum load set near room temperature is not quite optimal at high temperature. For example, we found that the optimum resistance of the 508 μm element increases by approximately 20% (7 mΩ) between room temperature and an applied temperature difference of 270 K. In principle, load optimization could be performed at every temperature step but that was not done here.

The maximum open circuit voltage measured for the n-type elements was 186 μV/K as shown in Figure 74. This is less that the 241 μV/K predicted from the bulk characterization measurements made by Marlow. There are a few possible reasons for this. We operated the element at elevated temperatures (beyond the range characterized by Marlow) and the Seebeck

coefficient of the material is known to decrease at high temperatures. We were also looking at thin elements, whose material properties may not match those of the much larger sample used by Marlow for their characterization measurements. Also, all of the temperatures measured between the copper blocks may not be dropping across the element itself. The electrical contact region of the element, the thermal interface between the element and the copper blocks, and the thermal spreading resistance into the copper block all contribute to the parasitic thermal impedance in series with the true thermal impedance of the element. A parasitic thermal interface resistance of 38 K/W/mm^2 at each contact is sufficient to explain the deviation of the measured Seebeck coefficient from the Marlow data, a very reasonable number.

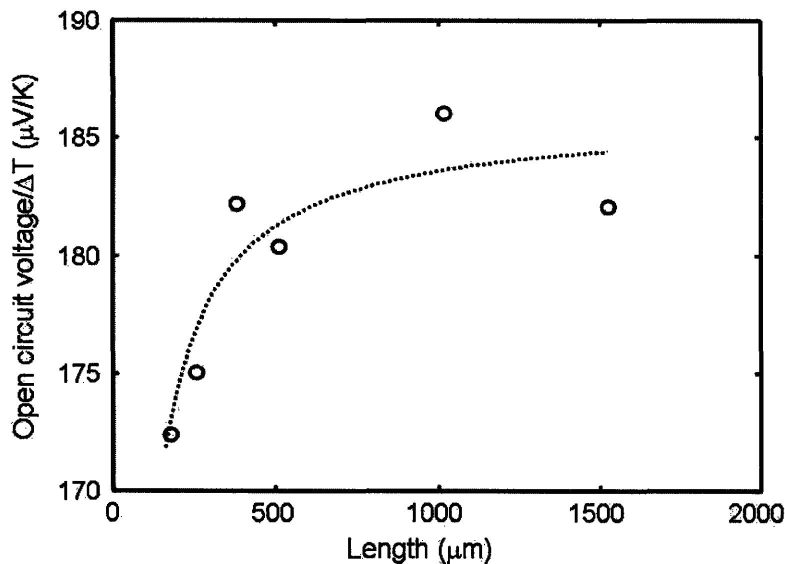


Figure 74. Open circuit voltage vs. length for Bi_2Te_3 elements.

Even without making reference to the Marlow Seebeck calibration we know that because there is no current flowing in this measurement the variations in the open circuit voltage with length are due to thermal parasitics or due to intrinsic variation in the materials in the different length legs. A best fit can be performed to the open circuit voltage using the ideal "lumped Seebeck" coefficient of the element in series with a fixed parasitic thermal resistance. When the elements are analyzed in this way (using a high temperature Bi_2Te_3 element thermal conductivity estimate of 0.024 W/cm^2/K) the decreasing voltages at short element lengths can be described by a parasitic thermal impedance of around 5.5 K/W/mm^2 in series with elements having a Seebeck coefficient of 186 μV/K (dotted line in Figure 74). If the Marlow thermal conductivity at 337 K is

used instead to estimate the material thermal conductivity, then the fit finds the same Seebeck coefficient and a parasitic series thermal impedance of 9.5 $K/W/mm^2$.
The thermal impedance is likely due to the finite interface resistance at the hot and cold side of the element. The 5.5 $K/W/mm^2$ parasitic interface resistance translates to an effective heat transfer coefficient of 18 $W/cm^2/K$ for the element while the 9.5 $K/W/mm^2$ parasitic interface resistance translates to an effective heat transfer coefficient of 10.5 $W/cm^2/K$ for the element. This is quite a bit larger than what is typical for most commercially packaged Bi_2Te_3 generators, but since the goal of this work is to study the power generation potential of single elements independent of more complicated packaging parasitics, the large heat transfer coefficient is ideal for us.

The power densities achieved in each of the devices at $\Delta T=200$ K is shown in Figure 75. At that temperature difference generated power densities tended to be around 1 W/cm^2 and dropped at both short and long lengths. The drop in power for the longer length elements is expected from the standard theory of how power density scales with length: as the element length increases its thermal resistance increases and the heat drawn through the device decreases, limiting the generated power. At short length the device's power generation is limited by the thermal parasitic resistance and the electrical parasitic resistance. The former was estimated using the 5.5 $K/W/mm^2$ parasitic thermal resistance calculation from earlier, but is insufficient to explain the drop. By assuming no variation in the intrinsic Seebeck voltages for the different length elements a line can be plotted using only the variation in measured optimum load with device length (dotted line in Figure 75). The similar shape in this curve to the measured power and its sharp drop at short length suggest that the existence of parasitic electrical resistance (that was measured in Figure 72) at short lengths explains the degradation of the power. The parasitic resistance in this setup is most likely not dominated by the electrical contact resistance of the elements but rather the wire leads and metal connections from the device to the active load.

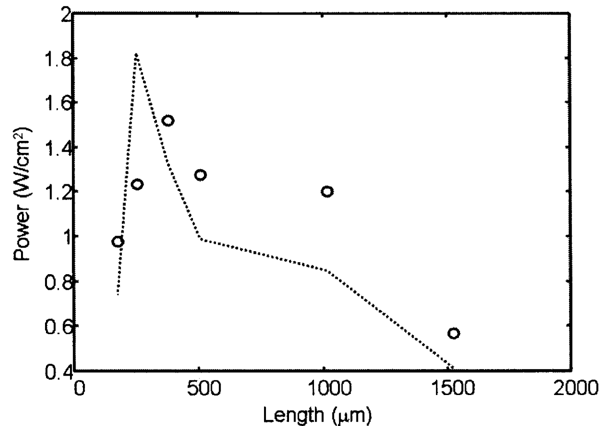


Figure 75. Generated power density vs. length for Bi_2Te_3 elements.

There are a few key conclusions that can be drawn from these Bi_2Te_3 measurements. Decreasing the element length appears to increase power until the thermal and electrical parasitic resistance become important, with the latter playing the more important role. The dependence of the generated power with length therefore shows a maximum consistent with the standard theory of TE element length optimization discussed in chapter 2. The weak dependence of the element open circuit voltages with length confirms that the setup is not dominated by thermal parasitics for these devices although they must be accounted for in order to accurately determine material parameters. It was possible to generate powers of over 20 mW (2 W/cm^2) from single elements.

4.4.2 Material parameter measurements of MAM element

In addition to the thinner samples discussed in the previous section we also obtained a 3.82 mm x 3.82 mm x 4.82 mm calibration sample from Marlow. Some low temperature material properties were measured by Marlow on a pellet of identical material (see Appendix 3). The ZT at the highest temperature tested by Marlow (337K) was 0.86. To test the performance of the high-temperature vacuum measurement system that we built, we measured the thermal conductivity and Seebeck coefficient of the calibration sample for high temperatures.

The sample was placed in the thermal test column so that thermal and electrical transport occurred in the direction of its long (4.82 mm) axis. Indium foil (thickness 50 μm) was used in addition to the contacts deposited by Marlow in order to ensure good thermal contact between the sample and the copper contacts in the test column. 795 psi of contact pressure was applied (measured when the device was cold), and a vacuum of 1.4 mTorr was drawn during the

measurements. The hot side of the generator was varied between 116 C and 330 C in 6 discrete steps while the cold side varied between 26 C and 41 C during the same steps. The temperature is assumed to drop linearly in the sample so that the average temperature in the sample is the mean of the hot and cold side temperatures. This mean temperature (together with the heat flux and open circuit voltage measurements) is used to plot an inferred thermal conductivity and Seebeck coefficient at temperatures between 69 C and 186 C. The results are shown below in Figure 76.

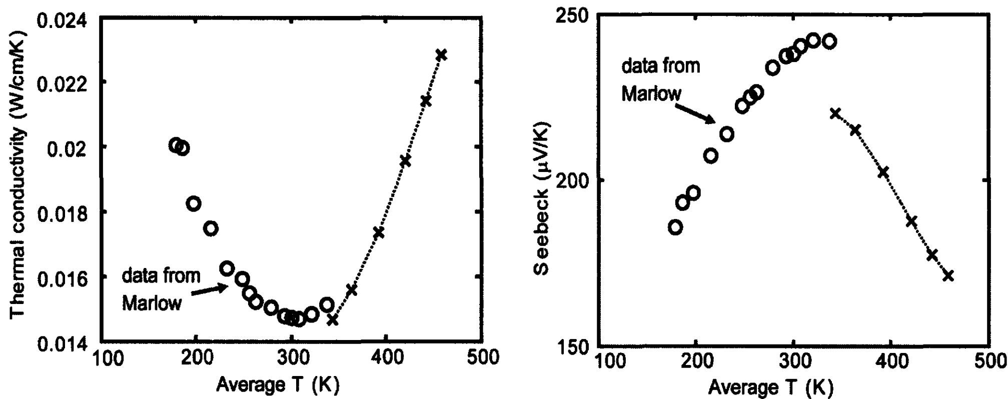


Figure 76. Thermal conductivity and Seebeck coefficient data from the Marlow MAM calibration sample. The circles are Marlow's measurements, and the x's are the measurements made using the high-temperature vacuum setup built for this thesis.

There are two important sources of error in these measurements. The first is that we are not measuring the material parameters at a precise temperature but rather over a range of temperatures that becomes larger the higher in temperature we go. The second is that we are most likely still experiencing a significant thermal interface resistance at the copper-indium-MAM interfaces. Based on our separate measurements of the thermal interface resistance at a copper-In foil-copper interface (shown earlier), we expect a thermal contact resistance of around 1 K cm^2/W and negligible thermal spreading resistance into the copper. Two such interfaces result in a total parasitic thermal impedance of 14 K/W compared with the 220 K/W thermal resistance expected from the element itself so that about 94% of the measured temperature is expected to drop across the element. This is not taken into account in Figure 76 and may help to explain the offsets between our measured data and the Marlow data (3% offset between the thermal conductivity datasets and 10% offset between the Seebeck coefficient datasets).

In light of the errors that may be due to thermal contact impedance it is worth discussing a couple of possible solutions. The most straightforward is simply to measure an element with a larger aspect ratio, i.e. an element whose length is longer or whose cross sectional area is smaller. The only problem with this approach (provided that the element can be made) is that the heat transfer measurement becomes more difficult. For the meter bar that we used for this measurement the thermal resistance of the element is so large relative to the meter bar that only a few degrees Kelvin are measured across the meter bar even at the largest element temperature differences. For longer element lengths we would very quickly be limited by the accuracy of the temperature difference in the meter bar. Redesigning the system with a longer (or thinner) meter bar might alleviate some of the problem at the cost an increased risk of parasitic convective and radiative heat loss from the sides of the meter bar, and longer measurement times (as per the results of Appendix 4).

Efficiency and power generation measurements were also made on the sample. However, because the resistance of the element (5 mΩ) was substantially smaller than our minimum load resistance in this measurement (27 mΩ) we were unable to impedance match so that both efficiency and power were substantially lower than their expected optimum values.

4.4.3 Generator performance from single elements

To test the extent to which generator performance could be judged from measurements of single elements, n-type and p-type elements were removed from a commercial HZ-2 generator and measured in the vacuum test chamber. The generator is shown in Figure 21. It consists of a checkerboard of 97 identical area square cross-section p-type and n-type elements 1.45 mm on a side. The length of elements was measured as 2.81 mm and the thickness of the module was 4.57 mm, with the difference between the two made up by the thickness of the electrical contacts on either side of the elements. The area between the elements in the checkerboard was filled by the electrically insulating Kapton "egg-crate" frame of the generator which has a book thermal conductivity value of 0.006 W/cm/K. The Kapton contributed a parallel "shunt" parasitic resistance to the model. The area ratio between the element and the checkerboard spacing gave the area ratio for heat spreading, which was used to calculate one of the series thermal impedances for the model.

The remaining ingredients needed for the model were the effective Seebeck coefficient, thermal conductivity, and electrical conductivity of the elements for an applied temperature difference of 200 C. The first two were measured in a manner identical to that discussed in Section 4.4.2, with two caveats. A parasitic contact resistance of 2 cm^2 K/W was assumed and subtracted from the elements thermal resistivity before calculating the material thermal conductivity and Seebeck coefficient. This parasitic turned out to make up 8% of the total thermal resistance of the device. Also, because of the long aspect ratio and high thermal resistivity of the elements and the high temperature difference used for the measurement, back of the envelope estimates revealed that the radiative heat transfer between the exposed faces of the copper contact plates would provide a small (but not negligible) fraction of the heat transport through the meter bar during the thermal conductivity measurement. To remove this contribution (~ 20% of the total heat transfer) a measurement with identical contact spacing but empty space between the contacts was performed so that the resulting parasitic heat flux could be subtracted from the element measurements. The cross-plane electrical resistance was too small to impedance match to but an estimate for the total internal resistance (the resistance of the element plus lead resistance) could be calculated from the difference between the open circuit and closed circuit voltage divided by the measured current. This resistance was used as the element resistance since we had no a priori estimate for the magnitude of the lead parasitic resistance compared to the element resistance. The power and efficiency of the elements were also measured.

The temperature drop across the TE elements was calculated using the ratio of the measured thermal conductivity of the elements to the calculated and measured parasitic series contact resistances. Using this temperature drop and the measured Seebeck coefficient, generated power could be estimated from the measured internal resistance of the elements (although the internal resistance included some lead parasitics). The total thermal resistance of the generator was calculated from the network consisting of the measured and calculated contact resistances, the measured element thermal resistances, and the calculated shunt thermal resistance of the Kapton egg crate. Knowledge of the heat transported through the generator and the power generated allowed the estimation of the module efficiency. Shown below in Figure 77 are the results of the parameter extraction. For further details the reader is referred to the documented MATLAB code in Appendix 6.

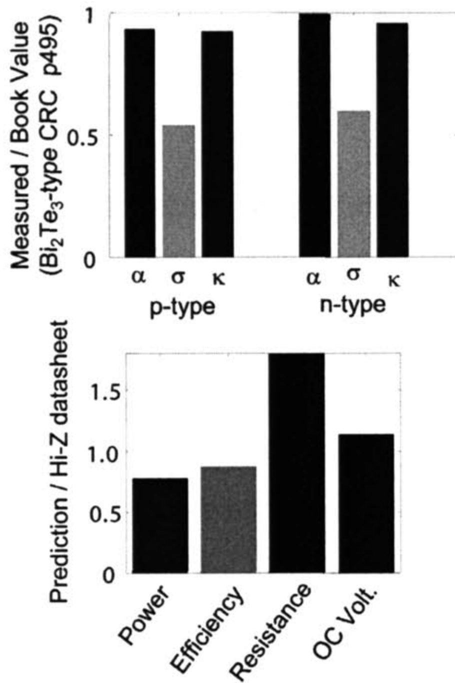


Figure 77. The top chart compares the measured element material parameters to book material values. The bottom chart compares the predicted performance of the generator based on the measured material against the stated performance of the generator from the manufacturer's data sheet.

The material parameters for the elements in the HZ-2 are not published, so we compared our measured performance to book values for Bi_2Te_3. The match is reasonably close except for our measured electrical conductivity which is half of what is expected. This is due to the presence of contact resistance of the leads, and serves to highlight the difficulty of cross-plane electrical conductivity measurements. The predicted generator performance is reasonably close to that predicted by the Hi-Z data sheet with the notable exception of the abnormally high resistance due again to the error in the cross-plane resistivity measurement. The decreased power and efficiency can most likely be traced to the same issue.

The results of the calculation provide some evidence that efficiency and power measurements on individual elements coupled with a very simple model for the parasitic thermal resistances in the generator allow the generator performance to be estimated. The main obstacle to higher accuracy is the cross-plane electrical resistivity measurement.

4.5 Thick film measurements

In this section we present thick-film measurements made on a lead-telluride based thermoelectric material from the Harman group at Lincoln Labs. By way of motivation, we first review their measurements of the material[21].

<u>The Harman n-QDSL measurements</u>

In September 2002, Harman, Taylor, Walsh, and LaForge published a measurement of room temperature thermoelectric figure of merit (ZT) in the range 1.3 to 1.6 using a novel material they termed a quantum dot superlattice[21]. Based on its citation record it appears to be one of the most important experimental breakthroughs in high-ZT nanostructured thermoelectrics.

The QDSL-A sample was grown in the PbSeTe/PbTe material system using a dedicated MBE reactor at Lincoln Labs. The idea behind the growth (as stated in the Science paper) is that a high density of PbSe nanodots are interspersed in a "three-dimensional slab matrix" of PbTe. These dots are intended to reduce the thermal conductivity of the material through selective scattering of phonons and also enhance the transport properties of the material by introducing a delta-function density of states and more favorable carrier scattering mechanisms.

Aside from their material properties and structure, the films grown by Harman et al. are remarkable from the perspective of metrology and thin film generator design for the thickness of their epitaxy (> 100 μm). Assuming a thermal conductivity of less than 0.010 W/cm^2/K, this thickness would give the elements a cross-plane thermal resistance of greater than 1.0 K cm^2/W. This is on the same order as the parasitic interface resistance that we measured for indium foil and is around 20 times larger than the thermal contact resistance measured for the liquid metal eutectic. It is also close to the thermal resistance we measured for pumped water heat sink, 1/0.66=1.5 K/W/cm^2. This is exciting, since it could mean that the material might be used efficiently in a traditional cross-plane geometry thermoelectric generator.

The idea behind the measurement of Harman et al. was to use the maximum cooling temperature difference of the device to set a lower bound on the ZT of their material. The known in-plane Seebeck and electrical conductivities are then used in conjunction with this minimum ZT to compute the maximum thermal possible thermal conductivity of the n-QDSL.

They begin by measuring the in-plane (transverse to the growth direction) Seebeck coefficients and electrical conductivities of two of the n-type QDSL samples, which we will term n-QDSL A and n-QDSL B1. The details of this measurement are not reported, but they find:

	In-plane Seebeck (μV/K)	q x mobility x carrier conc. = In-plane electrical conductivity (1/Ω/cm)
n-QDSL-A	-219	710
n-QDSL-B1	-208	530

In the text of the article, the electrical conductivity of QDSL-B1 is given as 585 1/Ω/cm. Next, two thermoelectric test couples are constructed. A sample from the same wafer as n-QDSL-B1 is attached to a gold ribbon (thickness: 25 μm, width: 250 μm, length: 5 mm) forming a thermoelectric couple, where the gold acts as the p-type leg. We denote this sample n-QDSL-B2 (thickness: 104 μm, width: 11 mm, length: 5 mm) to remind ourselves that it is actually a different piece of material from n-QDSL-B1. But by assuming it has identical material properties to n-QDSL-B1, we can summarize the thermoelectric properties and geometry of both legs of the couple in the table below:

	Area (thickness x width)	Length	α (μV/K)	σ (1/Ω/cm)	ρ (1/Ω/cm)	κ (W/cm/K)	Z_dT
n-QDSL-B1/B2	104 μm x 1.1 cm	0.5 cm	-208	585	1.71×10^{-3}	To be measured (0.0058)	0.40
Gold ribbon	25 μm x 250 μm	0.5 cm	2.9	4.17×10^5	2.4×10^{-6}	2.8	0.00

The thermal conductivity of the n-QDSL samples was the unknown part of the thermoelectric figure of merit, and the bulk of the paper describes their measurement of the thermal conductivity, from which a ZT can be calculated.

The next step in the measurement of the figure of merit of the n-QDSL sample was to apply a well-known formula for the maximum cooling that can be reached using a thermoelectric couple with the lumped figure of merit Z_d:

$$\Delta T_{max} = Z_d T_c^2 / 2 \quad (4.5.1)$$

Here ΔT_{max} is the temperature differential between room temperature and the maximum temperature that the thermoelectric couple can be cooled to, T_c. In this case the cooled junction is the gold-thermoelectric interface, which is adiabatically suspended in vacuum (although it is contacted by a thermocouple). The figure of merit Z_d in this case refers not to a single material, but to the lumped figure of merit for the thermoelectric couple, including also any parasitic series resistances and parasitic shunt thermal conductances. It could therefore be lower than the true material figure of merit, but not higher. The n-QDSL-B2 sample is measured by fixing one side of the gold and the thermoelectric at room temperature while passing current through the couple to cool the other junction, which is left (except for small parasitic resistances) adiabatic. The cold and hot side junctions are monitored with thermocouples, and for the optimum current, a maximum cooling temperature of 43.7 K was measured. According to Eq. 4.5.1 this means that Z_dT=0.4 at T=300 K. The Z_dT thus calculated can be related to the thermoelectric properties of the two materials in the thermoelectric couple according to the a simple expression given by Rowe in the introduction the CRC Handbook of Thermoelectrics [68]. (The expression can also be easily derived from the Onsager relations if the Thomson effect is neglected):

$$Z_d T = \frac{(\alpha_p - \alpha_n)^2 T}{KR}$$
$$R = \frac{L_p}{\sigma_p A_p} + \frac{L_n}{\sigma_n A_n} + R_{parasitic}$$
$$K = \frac{A_p \kappa_p}{L_p} + \frac{A_n \kappa_n}{L_n} + K_{parasitic} \quad (4.5.2)$$

This expression depends not only on the thermoelectric properties of the two materials in the couple, but also on the relative dimensions of the elements. It turns out that for two legs of known thermoelectric properties, there is a ratio between the aspect ratios of the n and p legs of the thermoelectric that maximizes the cooling:

$$\sqrt{\frac{\sigma_n \kappa_n}{\sigma_p \kappa_p}} = \frac{L_n A_p}{L_p A_n} \quad (4.5.3)$$

(For example, if the n-type leg of the couple was known to have a high thermal conductivity one could choose to use a longer n-type leg length to reduce the heat transport through that leg.)

When Eq. 4.5.3 is substituted into Eq. 4.5.2, the geometrical parameters can be eliminated, resulting in

$$Z_d T = \frac{(\alpha_p - \alpha_n)^2 T}{\left[\sqrt{\kappa_p/\sigma_p} + \sqrt{\kappa_n/\sigma_n}\right]^2} \tag{4.5.4}$$

This is the expression used in Eq. 3 of the paper by Harman et al. for the thermoelectric figure of merit of a TE couple. This is not correct usage of the formula, since he is using the formula to extract the ZT of a completely general (un-optimized) TE couple made of the n-QDSL material and a gold ribbon. (If the formula was correct, the formulas Eq. 4.5.4. and Eq. 4.5.2. would actually yield the same result, and they do not.) The error caused by this mistake is actually large, in part because of the vastly different thermoelectric properties of the gold compared to the n-QDSL material.

When the thermal conductivity is calculated from Harman's Eq. 4.5.3, he finds that the thermal conductivity of the n-QDSL is at most κ_n=0.0058. But when the thermal conductivity is calculated from the corrected expression for the couple figure of merit (Eq. 4.5.2), we find a different maximum: κ_n=0.00023. This is more than an order of magnitude below the limit predicted by the Wiedemann-Franz Law for metals or for semiconductors with typical band structures, to say nothing of the alloy limit in the solid. In any case, it is also more than an order of magnitude from what was claimed in the paper. Something is wrong.

One possible cause of the error is that a different sample was used for the in-plane Seebeck and in-plane electrical conductivity measurement (n-QDSL-B1) from the thermal conductivity measurements (n-QDSL-B2). We should also note that in support of the extraordinary ZT of their n-QDSL samples, Harman et al. attempt to compare their material to a conventional Bi_2Te_3-based material. Unfortunately, they use the same incorrect formula to evaluate the thermal conductivity, pointing out when they do so that the result of the calculation "does not agree with the κ value measured on the material at the University of Virginia." In light of what we have learned, this disagreement may not be so surprising. Unfortunately we cannot check whether the incorrect formula is again the cause of the error with the Bi_2Te_3-based material, because Harman et al. neglect to include the exact geometry of that sample or the gold ribbon

used with it as the p-leg of the couple in their paper. Another possible cause of the error could lie with three dimensional heat and current spreading effects. Figure 78 depicts the geometry (to scale) of the TE couple used for the measurement of Harman et al.

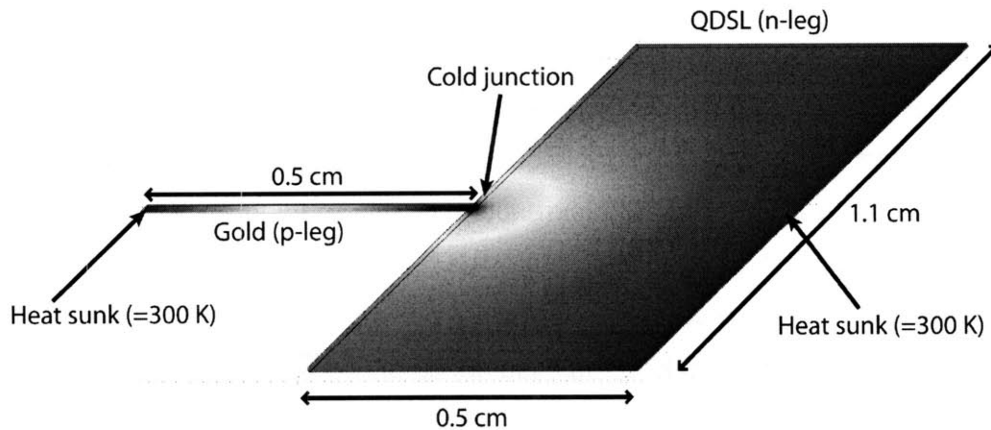


Figure 78. The geometry of gold-QDSL couple used by Harman et al.

The top (cold) junction of the thermoelectric couple is formed by a mating a gold ribbon that is 250 μm x 25 μm in cross section to a n-QDSL sample that is 11000 μm x 104 μm in cross section. The QDSL leg has an area 183 times bigger than that of the gold leg, and also a much lower thermal conductivity. Under these conditions, one would expect significant heat spreading in the QDSL due to the localized Peltier cooling at the junction between the gold and the QDSL (as shown in Figure 78). It seems unlikely that one may regard this as the simple 1-D heat and current transport assumed by the formulas the authors use to derive their ZT numbers. In fact, one would expect that an additional series thermal spreading resistance have to be included in the formula given by Eq. 4.5.2. This is a series thermal resistance, so it is not modeled by the $K_{parasitic}$ term in that equation. In fact, one would expect it has the effect of thermally isolating the cold junction beyond the amount predicted by Eq. 4.5.2, enhancing the measured ZT value. This effect (enhanced cooling due to heat spreading) has been experimentally observed elsewhere[99].

The remainder of their paper is devoted to an attempt to measure ZT a different way. Harman at al neglect Ohmic heating in their sample (n-QDSL-B2), and use their measured Seebeck coefficient from n-QDSL-B1 as a sort of thermometer. They infer the temperature drop across the couple due to the applied probe current from a measurement of the resulting (microvolt)

Seebeck voltage signal. They can then calculate the thermal conductivity of the couple from this temperature, and from exact knowledge of the geometry of both legs and the thermal conductivity of the gold ribbon they back out a thermal conductivity, this time equal to 0.0062 W/cm/K. This can be used with the measurement of electrical conductivity (from n-QDSL-B1 again) to calculate a ZT value.

In fact this is not very different from the earlier maximum cooling measurement. Only the range of the bias current was changed, and even less information from n-QDSL-B2 is actually being used in the calculation of ZT. The same issues with current and heat spreading at the cold junction are present, as is the issue with using different samples to measure different parameters and then combining the measurements to form ZT. While the Ohmic resistance of the couple is assumed to be around 9 mΩ (making the assumption of no Ohmic heating reasonable) the contact resistances getting into and out of the gold ribbon and the n-QDSL sample in the in-plane direction could easily be an order of magnitude greater, making the assumption of negligible Ohmic resistance questionable. In addition, the presence of any series thermal contact resistance between the QDSL material and the gold or between both materials and the heat sink will also cause errors. Neither of these parasitics are characterized in the paper.

The conclusion we draw from this analysis is that more measurements are needed to be quantify the ZT of the n-QDSL material. A measurement should be on the same sample with the properties being measured at the same time, if possible. The heat and current transport should be kept one-dimensional so that we can interpret the results easily, and we would prefer not to have to make assumptions about the parasitic thermal and electrical contact resistances. This is exactly what the high-temperature vacuum test bench is designed for, and we turn next to those measurements.

Our n-QDSL (G-207) measurements

We measured the Seebeck coefficient and thermal conductivity of a sample of the n-QDSL material (also called G-207) in the vacuum test chamber, using similar methods to those described in Section 4.4.2. The sample was 95 μm in thickness (measured by Lincoln labs), and 0.04 +/- 0.005 cm^2 in area (measured by us). The 0.317 cm^2 round copper meter bar and contacts were used for the measurements.

The sample was extremely fragile, so that even the unsupported weight of the heating fixture could crush the material. The issue was resolved by hanging the heater from springs and monitoring contact with a pressure gauge capable of gram accuracy (Omegadyne strain gauge from Omega). Aside from the issue of contact pressure, a challenging step in mounting the sample was applying enough of the GaSn liquid metal eutectic (supplied for this purpose by Bob Reeder of Lincoln Labs) to each side of the sample to facilitate electrical contact while not allowing it to well up around the outside of the sample and short the top copper electrode to the bottom copper electrode. By polishing the copper contacts and heating them it was possible to get the GaSn to wet the copper, making good contacts possible.

A vacuum of 1.6 mTorr was used to eliminate unwanted convective heat transfer and approximately 30 grams of contact pressure was used. Only two temperature set points were measured due to concerns regarding sample degradation. Once the temperatures were reached and the meter bar temperature stabilized 7 measurements (over 24 seconds of time) were taken at the first temperature and 45 measurements were taken at the second point (over 180 seconds of time). The average of each of the results is summarized in the table below for the two sampled temperatures.

	ΔT (K)	T_{hot} (C)	T_{cold} (C)	κ (W/cm/K)	α (μV/K)	Q_{out} (W)
Measurement 1	66.69	108	41	0.0088	-216	2.48
Measurement 2	80.47	125	45	0.0085	-221	2.89

Here ΔT is the temperature drop across the element, T_{hot} and T_{cold} are the corresponding hot and cold side temperatures (rounded off to the nearest degree), κ is the thermal conductivity, α is the Seebeck coefficient, and Q_{out} is the heat through the meter bar below the sample.

There are two sources of error with which we must concern ourselves. Since we are not discussing closed circuit measurements, we need not concern ourselves with electrical parasitic resistance. The thermal parasitic contact resistance between the copper contacts and the TE element are a possible source of concern. However, elsewhere (Section 4.3.3) we measured the contact resistance at a copper-liquid metal-copper interface as 0.043 K cm^2/W. Because the equivalent area thermal resistance of the element is L/κ=1.91 K cm^2/W, we estimate our error is

roughly 8% from the two interface thermal resistances; i.e., the material thermal resistance we measured is 8% larger than the true resistance due to this effect. An opposite parasitic effect is a possible cause for concern as well. The open area of the hot copper contact is presumably radiating heat towards the cold side contact according to the Stephan-Boltzmann law. If we conservatively estimate that the copper has an emissivity of 0.9 (in reality it was probably much lower) we can calculate that the radiative heat was at most 0.046 W. This is small compared to the conductive heat we measured (~ 2.5 W). The total error from spurious radiative heat transfer is estimated at under 2%; i.e., the material thermal resistance we measured is 2% smaller than the true resistance due to this effect. We think the main error in the Seebeck coefficient measurement was due to the thermal contact resistance, so that the true material Seebeck might be 8% larger than the one we measured. The main error in the thermal conductivity was actually probably due to our uncertainty in the element area. As a result, we estimate our combined error in the thermal conductivity measurement at around 15%.

After this measurement was taken, the vacuum chamber was allowed to fill with air at standard pressure and a long (> 24 hr) measurement was taken (Figure 79). (Air was allowed into the chamber because it was felt that the device would survive for longer with some partial gas pressure opposing a suspected out-gassing degradation process). When air was allowed into the chamber, the measured Seebeck increased to 232 μV/K. We speculate this might be due to improved thermal contact resistance with the air present. The measured thermal conductivity of the device also increased, due to the parasitic convection that is possible between the hot and cold copper contacts in ambient conditions. The device maintained its operation for roughly 2 hours, at which point the Seebeck coefficient suddenly began to steadily decrease. At 20 hours after the start of the measurement, we saw a new degradation mechanism as the thermal conductivity of the sample began to steadily increase, changing by a factor of two over the next 10 hours. The exact causes of these effects are still not known.

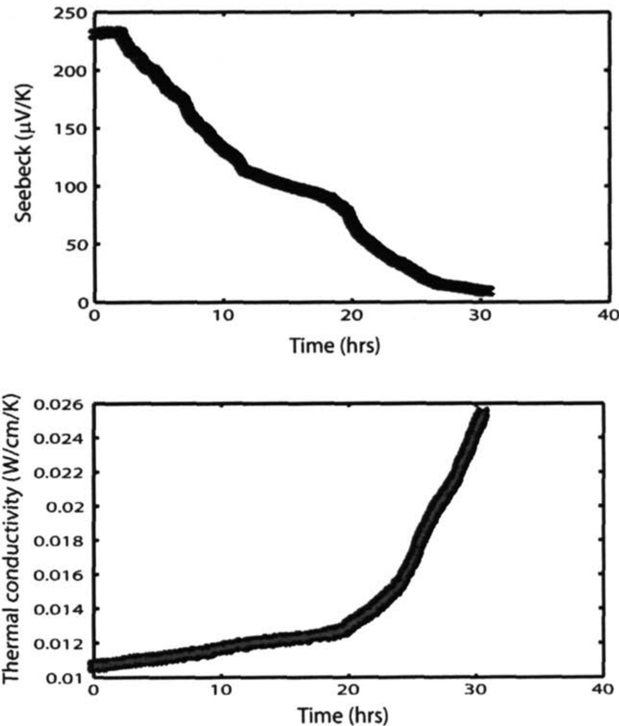


Figure 79. A > 1 day-long measurment of the Seebeck coefficient and the thermal conductivity of the Lincoln Labs G-207 (n-QDSL) sample. This long measurement was done in air to prolong the life of the device.

In conclusion, we simultaneously measured the cross-plane thermal conductivity and cross-plane Seebeck coefficient of the Lincoln Labs G-207 material (also known as the n-QDSL material). At an average temperature of 358 degrees K we measured α=-221 +/- 18 μV/K and κ=0.0085 +/- 0.0013 W/cm/K. These results differed significantly from those of Harman et al. (α=-208 μV/K and κ=0.0058), and we have explored some possible reasons why this might have been so. We were unable to measure the cross-plane electrical conductivity of the sample, and so we will not estimate the ZT.

4.6 Thin film measurements - ErAs nanodot superlattice

Recently, significant improvements have been made in the material parameters of thermoelectric (TE) elements using nanostructured growth techniques. Heterostructures, quantum confinement, and embedded nanoparticles have been suggested as a means to substantially increase the Seebeck coefficient (α) and electrical conductivity (σ) and decrease

the thermal conductivity (κ) relative to that of bulk materials[21-23]. The maximum generated power density from a TE material is the other important quantity for applications and to first order it depends in a simple way on σ and α (Eq. 2.2.4).

The performance of thin-film devices under strongly non-equilibrium conditions (large temperature gradients, large generated current densities) has proven difficult to measure, particularly in the cross-plane (growth) direction. This is in part a general challenge of thin-film thermoelectric property characterization. While cross-plane thin-film thermal conductivities can be obtained using the 3-ω technique[57,100] the measurement of cross-plane Seebeck coefficients and electrical conductivities is somewhat more problematic due to the small geometry and the large influence of electrical parasitics, although techniques based on parameter fitting have been attempted[87,88,101,102]. For power generation (non-equilibrium) measurements, further difficulties arise. The load required to generate significant current in a single element is typically very small, on the order of mΩ. Also, many TE thin-films are grown on substrates and separating the effects of the thin-film from those of the substrate is difficult. Maintaining a large temperature difference across a thin-film in the cross-plane direction requires that a correspondingly large heat flux be removed at the cold side of the element. Single element measurements in strongly non-equilibrium conditions can provide information unavailable from wafer-level measurements of α,σ,and κ. The spatial and temperature variation of the material parameters within the superlattice can have large effects on performance, as can electrical parasitic resistances associated with transport through the multilayer structure. The measurements presented here allow nanostructured materials to be tested under conditions similar to those expected in a real power generation application.

The samples consist of a superlattice of self-assembled semi-metallic ErAs dots in a matrix of $In_{0.53}Ga_{0.47}As$, with InGaAlAs barriers for enhanced ZT. Data on non-superlattice samples of the same dots have previously been published in Zide et al.[103] (the material is designated at Structure B of Fig. 1 in that paper). Figure 80, reproduced from the same paper, shows a transmission electron microscope image of the dot material. The ErAs dots and the superlattice both scatter phonons participating in cross-plane heat transport, reducing κ. The free electron concentration is increased by the ErAs dots[104], increasing σ. The superlattice contains 160 periods, and each period consists of 10 nm $In_{0.52}Ga_{0.29}Al_{0.19}As$ barriers (grown as a digital alloy of InGaAs and InAlAs [both lattice matched to InP]) followed by 20 nm of InGaAs with 0.3%

ErAs and co-doped with Si at $8x10^{18}cm^{-3}$. The total superlattice length is approximately 5 μm. Each superlattice is grown at 763 K on a buffer layer consisting of 150 nm of n-InGaAs and 20 nm n-InGaAs on an n-InP substrate. All samples were grown by molecular beam epitaxy. Four substrate thicknesses were studied: 474 μm, 353 μm, 246 μm, and 180 μm. Electrical contacts were made using Ni/GeAu/Ni/Au stacks, and a 5 μm indium layer was deposited on both sides to aid the thermal and electrical contacts.

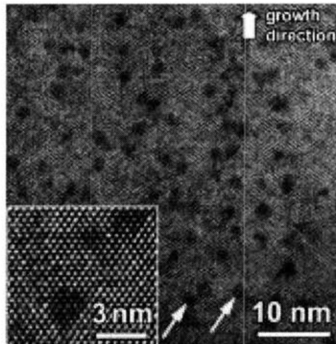


Figure 80. Cross-sectional transmission electron micrograph of ErAs nanodot material, confirming the randomly spaced nanometer-sized ErAs island formation.

As discussed in Chapter 2, the power generated in a load R_L from a uniform TE element with an applied temperature difference ΔT (Eq. 2.2.4) can be rewritten

$$P = I^2 R_L = \left(\frac{\alpha \Delta T}{R_{int} + R_L} \right)^2 R_L, \qquad (4.5.1)$$

where R_{int} and α are the internal resistance and Seebeck coefficient of the element, respectively. The generated power depends quadratically on the applied temperature, highlighting the importance of a large temperature gradient for power generation. If the expression for power in Eq. 1 is optimized with respect to the load (assuming that ΔT is somehow held fixed), one finds that the optimum load occurs at $R_L = R_{int}$. As with many other thin-film thermoelectrics, the elements studied here have R_{int} on the order of mΩ due to the high electrical conductivity of the thin-film + substrate and the short aspect ratio of thin-film elements (area=0.01 cm^2 for all samples).

The test setup described previously in this section was used to apply a temperature difference to a single element and measure the resulting power in a load at or near the optimum for power

generation. As before, one of the main difficulties lies in ensuring that most of the applied temperature difference is dropped in the elements rather than the surrounding blocks. Finite element modeling of the heat spreading in the experimental setup indicates that approximately 82% of the measured temperature difference was applied across the thickest element (474 μm), and approximately 74% of the measured temperature difference was applied across the thinnest element (180 μm). An additional fraction of the remaining temperature difference was most likely lost to thermal contact resistance. Assuming the total interface thermal resistance was similar to the resistance found in the Marlow element study earlier in this chapter (5.5 $K/W/mm^2$), the percentage of the remaining temperature expected to be lost to the interfaces is between 40% (for the thickest element) and 58% (for the thinnest element).

Only a fraction of the (remaining) temperature drop across the element will be across the superlattice, which is the material of most interest to us. The table below summarizes how the temperature and the resulting open circuit voltages expected from the superlattice and the substrate are expected to break down between the superlattice and the substrate (assuming the thermal conductivity of the superlattice is 0.036 W/cm/K and the thermal conductivity of the substrate is 0.68 W/cm/K).

	Substrate thickness			
	474 μm	353 μm	246 μm	180 μm
% of the temperature drop expected across superlattice	16%	20%	26%	33%
% of the open circuit voltage expected from superlattice	40%	47%	56%	64%

The end result is that only around 7.8% (for the thickest element) to 10.2% (for the thinnest element) of the measured temperature difference will actually be dropped across the superlattice.

The power is measured using the active load described earlier. An example of the power optimization procedure is shown in Figure 81 for the 474 μm device with a 154 K applied temperature difference. Near a load of approximately 20 mΩ, a maximum in the power is

obtained. The minimum load that could be applied using this setup was approximately 15 mΩ, limited roughly equally by the FETs residual impedance, the 5 mΩ sense resistor, and wiring parasitics.

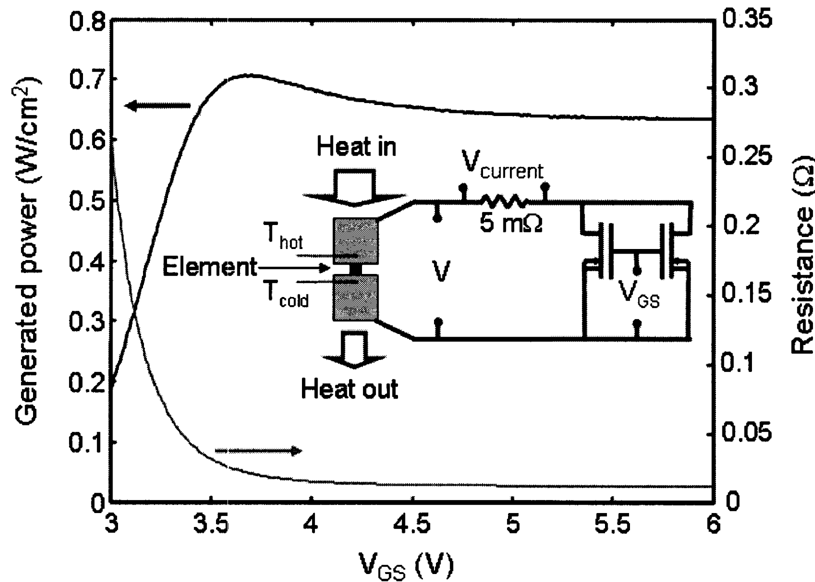


Figure 81. Load and generated power vs. active load control voltage. Inset shows schematic of test setup.

During a measurement, the device's impedance was optimized (for the 474 μm substrate device) or nearly optimized (for the other lengths) at elevated temperature, and then the device was allowed to cool while the six measurements were made at regular time intervals: hot side temperature, cold side temperature, open circuit voltage and current, and closed circuit voltage and current. Shown in Figure 82 are the resulting power densities as a function of the applied temperature difference for each substrate thickness.

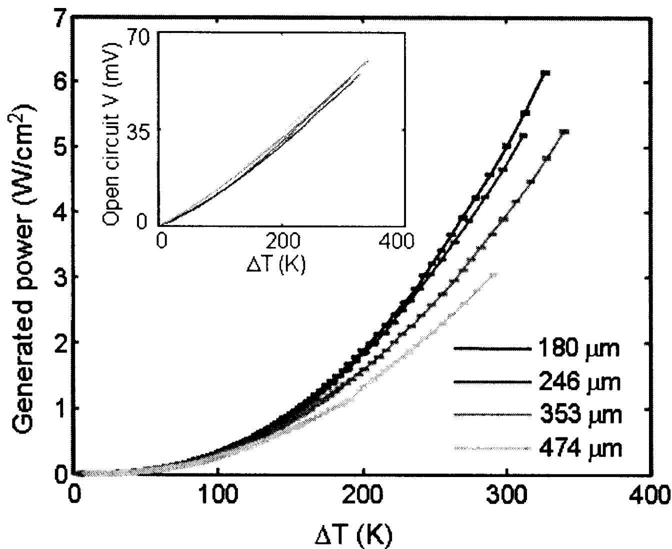


Figure 82. Generated power densities for ErAs superlattice material on differing substrate thicknesses, as a function of applied temperature difference. Inset shows open circuit voltages over same temperature range.

The curves follow the parabolic profile predicted by Eq. 4.5.1. There is also a trend towards higher power as the length of the substrate decreases. This occurs despite that fact that for the shortest three elements the device resistance was too small to optimally match. There are two ways to understand the increase in power with shorter substrate. First, for shorter substrates the internal resistance (R_{int}) of the element is smaller, resulting in a smaller denominator in Eq. 1 and generating more power. This is a well-known benefit of using shorter TE elements. Second, a shorter substrate thickness results in a larger fraction of the temperature drop occurring across the superlattice. Because the superlattice is a much better TE material than the substrate, this results in more generated power. The open circuit voltage of the four samples decreases slightly (~10%) from the shortest to the longest length, suggesting that the first effect is the dominant one. Using values for the Seebeck coefficients and thermal conductivities of the substrate and superlattice at room temperature, the expected open circuit voltage near room temperature for an applied temperature difference of 50 K can be calculated for each sample length and compared with the measured open circuit voltages (Figure 83). The necessary parameters for calculating the expected open circuit voltages were the Seebeck coefficient of the superlattice[102] (290 μV/K), the thermal conductivity of the superlattice[105] (3.6 W/m/K), the Seebeck coefficient of the substrate[106] (80 μV/K), and the thermal conductivity of the substrate[107] (68 W/m/K). The dominant error between the expected and measured voltages is likely due to variations in thermal contact resistance between the devices.

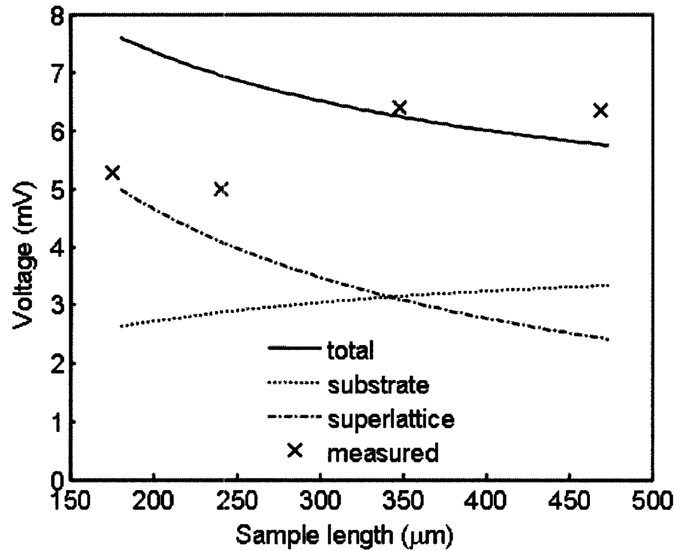


Figure 83. Measured and expected open circuit voltages for different substrate lengths. The expected contribution from the superlattice and the substrate are shown.

Because the open circuit voltage is measured through the range of applied temperatures, one can measure the lumped Seebeck coefficient of the material by fitting a smooth curve to the open circuit voltage and taking its derivative with respect to temperature. The results of such a procedure using parabolic fits are shown in Figure 84. Also shown is the analogous curve (dotted line) for a 380 μm-thick sample of n-type bulk Bi_2Te_3 [108] optimized for room temperature. All of the superlattice samples show increasing open circuit voltage with increasing hot side temperature, while the Bi_2Te_3 performance drops at higher temperatures. The degradation of the Seebeck coefficient at higher temperatures is well documented in bulk samples of Bi_2Te_3, resulting from the onset of intrinsic behavior due to the relatively small 0.2 eV bandgap of the material [109]. The fact that the Seebeck coefficient of the superlattice samples increases with temperature even at temperatures approaching the growth temperature of the superlattice is reassuring from the standpoint of device stability, and suggests that enhancements in performance may be possible by operating the superlattice devices at higher temperatures.

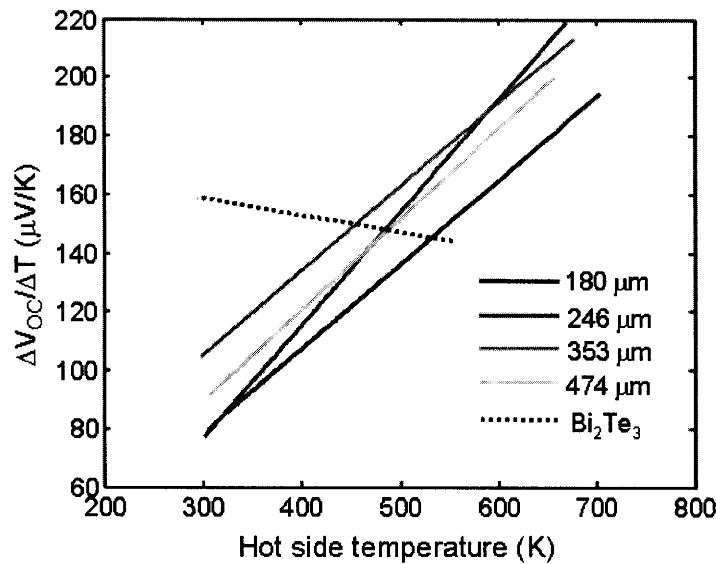


Figure 84. Slope of open circuit voltages with respect to the hot side temperature for ErAs superlattice material and for a representative Bi_2Te_3 sample.

In summary, the performance of single thermoelectric elements made from semimetallic ErAs nanoparticles embedded in a semiconducting InGaAs matrix with InGaAlAs barriers was measured under large temperature gradients for identical thin-films on top of differing substrate thicknesses. Power densities of greater than 6 W/cm^2 were measured at or near the optimum load condition. The generated power displays an increasing trend with decreasing substrate thickness, consistent with the standard model, and the slope of the open circuit voltage with temperature increases with temperature for all of the samples, suggesting that further performance benefits may be possible for higher operation temperatures. Unfortunately, it is also clear from this analysis that the uncertainties in all of the additional parasitic thermal and electrical impedances in the measurement make the accurate determination of material parameters for these films very difficult.

One other issue is worthy of clarification. The power density generated in the 180 μm substrate ErAs nanodot element is 6 W/cm^2, which is three times that of the MAM Bi_2Te_3-based element measured in Figure 73. The question arises: why is this so, and what (if any) implication does it have for the relative ZT of the materials? The question can be resolved by first noticing that at the highest power levels in each case, both elements had roughly equal open circuit voltages. (When the measured Seebeck coefficient of the MAM material (185 μV/K) is multiplied by the temperature difference (275 C), we find an open circuit voltage of 50 mV for the element; when

the Seebeck coefficients for the ErAs nanodot material (280 μV/K) and the InP substrate (80 μV/K) are used in combination with the fact that roughly a third of the temperature drop in the composite element is expected across the ErAs nanodot material and 2 thirds across the substrate, we find an open circuit voltage drop of 48 mV.) Since the generated power density is a function of only the Seebeck voltage and electrical resistances, the difference in power is presumably due to the electrical impedances. If the cross plane electrical resistance of the ErAs nanodot element (and its associated lead parasitic resistance) were 1/3 of the cross plane electrical resistance of the Marlow material (and its associated lead parasitic resistance), then the difference in generated power has been explained. This does not seem unreasonable given the fact that the ErAs nanodot element is about a third the thickness of the MAM element, although we do not have cross-plane electrical conductivities for the ErAs nanodot element (the macroscopic MAM sample has a conductivity of 650 1/Ω/cm). The different electrical parasitic contact resistances due to the indium on the ErAs nanodot samples vs. the dry contacts to the Bi_2Te_3 may also play a role. Nevertheless, the higher power density does not translate to higher ZT, because it does not factor in the huge difference in thermal conductivity between the materials. On a per-length basis, the MAM element is around 30 times less thermally conductive than the ErAs-nanodot+substrate composite, largely because of the relatively high thermal conductivity of the InP substrate (0.68 W/cm/K at 300K).

The effects of parasitic contact resistance are not confined to single element characterization. The performance of a generator design using the ErAs nanodot material (Gehong Zeng, Bowers Group, UCSB) was measured in our high temperature vacuum test station, and suffered from many of the same issues as the individual elements. The generator was made with a flip chip bonding process out of 400 elements (12 μm thick), 200 n-type elements and 200 p-type elements. Electrical contacts were made between the elements using a gold-tin alloy, and mechanical stability and electrical insulation were provided by building the generator between 140 μm thick AlN plates. A photograph of the generator and a scale drawing of a single element is shown in Figure 85.

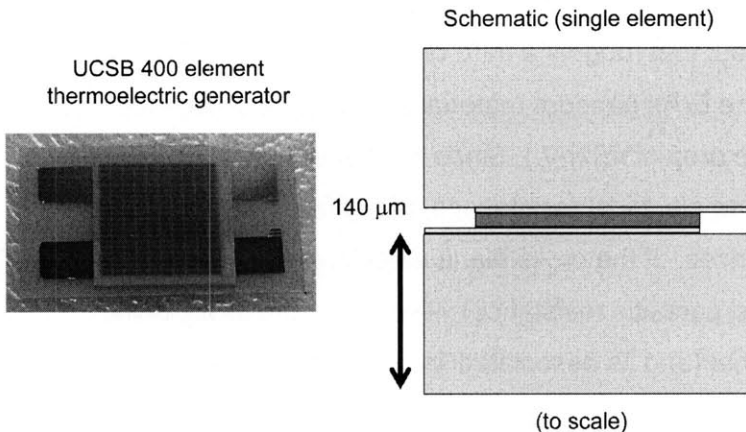


Figure 85. UCSB 400 element thermoelectric generator (photo courtesy of Gehong Zeng).

The total thermal resistance of the generator (area=0.36 cm^2) was measured to be between 2.55 K/W and 3.1 K/W between 26 C and 70 C, indicating that a heat sink with a heat transfer coefficient of around 1 $W/cm^2/K$ is needed in order for 50% of the temperature drop to occur across the generator. Much of the remaining temperature is likely dropping across the thermal contacts to the device (indium foil), based on the estimated thermal resistance of the indium contacts (1 K cm^2/W). There is also significant a temperature drop across the 140 μm thick AlN plates, which we compute will be roughly 40% of the total temperature drop applied to the generator. Even more serious is the parasitic electrical resistance for this design: the total resistance for the elements (assuming a conductivity of 1000 1/Ω/cm and uniform 1-D current density) neglecting parasitics ought to be on the order of 1.5 Ω, but the measured resistance of the generator was 152 Ω. This decreases the power by an additional factor of 100. Our measurements confirmed this picture, with an applied temperature difference of 32 C (the mean element temperature was estimated to be 70 C), we measured 0.05 mW of generated power for a heat flux through the generator equal to 12.89 W. The efficiency is therefore 0.0004%, more than 3 orders of magnitude below that attainable using a standard commercial generator for the same temperature drop, and in rough agreement with the values for the electrical and thermal parasitics we have estimated. The main point to take from this design example is the crucial importance of issues beyond the material ZT in determining power generation performance. Recent improvements to the design of this generator have addressed some of these issues, although not to the point of making the device competitive with commercial technologies.

4.7 Conclusions

In this chapter, we designed and tested general tools for thermoelectric real-conditions device testing. Using these tools, we measured thermal conductivities, Seebeck coefficients, power, and efficiency in elements ranging from bulk commercial elements to epitaxial thin-films. As expected, the measurements become progressively more difficult for the thinner elements as we eventually arrive at the limits of applicability of macroscopic heat and electrical measurements. We found that for elements whose electrical resistance was greater than 15 mΩ it was possible to electrically impedance match to the element using an active load and operate the element with efficiencies and power densities comparable to a generator built from the element. Likewise, for elements with thermal resistances greater than the thermal contact resistance of the interface it was possible to measure material thermal conductivities and open circuit (Seebeck) voltages with confidence as well as to predict generator performance.

The deleterious effects of the parasitics are a serious issue for packaged generators built from thin-films as well as for characterization of the films themselves. We believe that as a result, there are serious unresolved issues with many of the thin-film thermoelectric measurements in the literature, including the result of ZT>2 using the transient Harman technique[22] and the result of ZT>1.3 using the maximum cooling temperature (and related techniques)[21]. We have pointed out that the latter result appears to be based in part on an incorrect usage of a formula for the thermoelectric figure of merit.

Section 5: Nanoscale thermal imaging

5.1 Overview

This chapter presents thermal measurements made using the lock-in CCD thermoreflectance technique. Section 5.2 reviews the physics responsible for the change in the reflectance of a solid due to a change in temperature. Section 5.3 discusses thermoreflectance imaging in the context of other related thermal metrology techniques. Section 5.4 presents our theory of the thermoreflectance measurement process, in which fundamental sources of noise are propagated through the signal flow of the measurement in order to develop precise estimates of the measurement error. Section 5.5 experimentally confirms this theory. Section 5.6 presents a method for calibration of the thermoreflectance coefficients using a microthermocouple measurement, and checks the calibration against an independent thermal measurement using the 3ω technique. A fundamental limit to attainable resolution is recognized in the CCD camera's A-to-D converter nonlinearity. Section 5.7 presents a series of measurements on polysilicon gate FET transistors designed to test the range of the measurement. Section 5.8 discusses the measurement of the thermoreflectance of a sample of the Bi_2Te_3-based MAM material in the cross-plane (relative to thermal and electrical transport) direction, and attempts to model the dynamic temperature profile.

5.2 Physics of thermoreflectance

For much of the subsequent work in this thesis, a highly phenomenological understanding of the thermoreflectance phenomenon is sufficient to understand and interpret the results. However, it is helpful to have a basic picture of the physical origin of the thermoreflectance signal both to understand possible limits of the validity of this phenomenological treatment and to be able to predict, to some degree, the energies of features in the thermoreflectance spectra of materials. Thermoreflectance is a form of modulation spectroscopy whereby the complex dielectric response function $\varepsilon(E = \hbar\omega) = \varepsilon_1 + i\varepsilon_2$ of a solid material is probed while undergoing temperature modulation[110,111]. For a semiconductor near the band-edge, the absorption term ε_2 can be related to the band structure through the expression

$$\varepsilon_2(E) = \frac{4e^2\hbar^2}{\pi\mu^2E^2}\int dk|P_{cv}(k)|^2\delta[E_c(k) - E_v(k) - E]. \tag{5.2.1}$$

Here E_c and E_v refer to the conduction and valence bands, respectively, and P_{cv} is the momentum matrix element describing the optical coupling between them. If ε_2 is known, ε_1 can be calculated by using the Kramers-Krönig relation. The complex index of refraction is related to ε_1 and ε_2 according to $\tilde{n} = \sqrt{\mu(\varepsilon_1 + i\varepsilon_2)}$, and the reflectance is related to the complex index according to $R = \left|\frac{1-\tilde{n}}{1+\tilde{n}}\right|^2$.

The measured fractional change in reflectance due to a general external modulation can be expressed as $\frac{\Delta R}{R} = \alpha(\varepsilon_1,\varepsilon_2)\Delta\varepsilon_1 + \beta(\varepsilon_1,\varepsilon_2)\Delta\varepsilon_2$, where $\alpha(\varepsilon_1,\varepsilon_2)$ and $\beta(\varepsilon_1,\varepsilon_2)$can be easily derived from the expressions above[112]. For the case of an applied temperature modulation (e.g. using a small heater below the surface of the sample), a phenomenological model is typically used to describe the effects of the temperature on $\Delta\varepsilon = \Delta\varepsilon_1 + i\Delta\varepsilon_2$. The change in the complex dielectric function near a semiconductor critical point is understood to be caused by a shift in the bandgap energy E_g [113] and by a change in the phenomenological broadening parameter Γ:

$$\Delta\varepsilon = \frac{d\varepsilon}{dE_g}\frac{dE_g}{dT}\Delta T + \frac{d\varepsilon}{d\Gamma}\frac{d\Gamma}{dT}\Delta T \quad . \qquad (5.2.2)$$

The shift in bandgap energy due to changes in temperature may be partly understood from $k \cdot p$ theory. As the temperature increases, the lattice constant of the crystal increases, the band gap decreases accordingly. It turns out that this effect is usually actually smaller than a more subtle shift in the energy gap owing to the electron-phonon interaction. This may be understood in a qualitative sense by considering the change occurring in the energy spectrum (band structure) of the semiconductor as the temperature is increased from absolute zero up to room temperature. The allowed energy states in the semiconductor bands at zero temperature have very well-defined, precise energies. As the temperature increases, the states energetically couple to the bath of non-zero temperature phonons present in the crystal. The coupling with the phonon reservoir introduces an uncertainly in the energy states (the fluctuation-dissipation relation is one way to see this). The effect of this uncertainty on the gap energy can be understood by visualizing the convolution of a broadening line-shape (e.g. a Lorentzian) with the bare energy levels predicted by $k \cdot p$ theory. The "smearing" due to the convolution causes the bandgap to decrease. This shift in can be rigorously calculated using perturbation theory (e.g. the so called Debye-Waller and self-energy terms).

The physics responsible for the broadening parameter Γ is less clearly discussed in the literature, although it was originally introduced to capture the large excitonic effects present at some semiconductor critical points[114]. As far as we know, it is a purely phenomenological term measured through spectral curve fits[115]; we are not aware of any published comparisons between ab initio calculations of Γ and measured thermoreflectance spectra.

Depending on the dimensionality of the critical point of interest, detailed analytic expressions for the terms in $\frac{d\varepsilon}{dE_g}$ and $\frac{d\varepsilon}{d\Gamma}$ can be used to fit thermoreflectance data to theory while extracting the critical point energies[116,117]. An example of this from Miyazaki et al.[118] is shown in Figure 86. The model of Eq. 5.2.2 has also been extended with some success to fit excitonic lines near the band-edge[111]. If the energy-gap shift of the first term in Eq. 5.2.2 is assumed to be dominant (often asserted but not always true), the thermoreflectance spectrum has a simple interpretation, it is proportional to the first derivative with respect to optical frequency of the joint electronic density of states. For this reason, it is referred to as a first-order modulation spectroscopy (other forms of modulation spectroscopy include piezoreflectance (1st order) and electroreflectance (3rd order)). The heightened sensitivity of these derivative techniques near critical points made them a popular tool for determining the energies of semiconductor and metallic critical points[112,117,119].

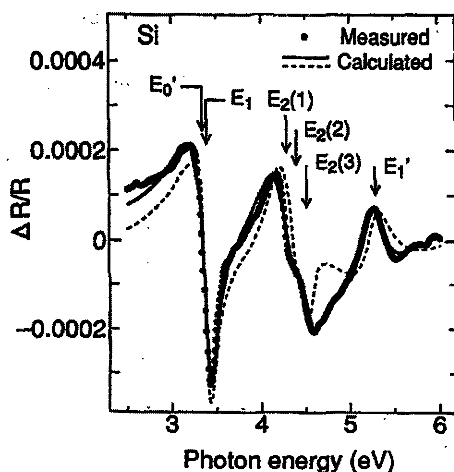


Figure 86. Thermoreflectance spectrum of silicon. (Miyazaki, 1995)

In addition to the basic study of the band structure and dielectric response functions of semiconductors, the thermoreflectance phenomenon has also been employed in a pump-probe configuration to study the picosecond time-scale thermal relaxation of carriers and phonon transport[120]. It has been applied in both the time-domain ("transient thermoreflectance method" [121]) and the frequency domain ("photomodulated thermoreflectance"[122,123]) to characterize thin-film thermal diffusivities and other thermal properties.

While the thermoreflectance phenomenon is an interesting topic in its own right, for the work in this thesis, thermoreflectance is used simply as a thermometer[124], and not as a probe of the material property. The quantity of interest is the thermoreflectance coefficient, with all of the physics causing a material's reflectance modulation lumped into one parameter:

$$\Delta T = \left(\frac{1}{R}\frac{\partial R}{\partial T}\right)^{-1}\frac{\Delta R}{R} = \kappa_{TR}^{-1}(\lambda)\frac{\Delta R}{R}. \qquad (5.2.3)$$

Although a measurement of the temperature need rely on only one wavelength, knowledge of the entire spectrum $\kappa_{TR}^{-1}(\lambda)$ is useful since an excitation wavelength then can be chosen to maximize the sensitivity of the measurement[125]. For most metals and semiconductors, the thermoreflectance coefficient is in the neighborhood of $\kappa_{TR} \sim 10^{-5}$ 1/K to 10^{-4} 1/K, meaning that the reflectance changes only by one part in 10^5 or 10^4 per degree Kelvin of temperature change. Thus a highly sensitive measurement is required to obtain quantitative information about the temperature of the device from the technique.

It is important to note that the thermoreflectance spectrum can be sample geometry dependent. Crystal orientation and incident light polarization are of obvious importance to the sensed dielectric functions. For thin-films whose thickness is on the order of the wavelength of the incident light (or samples encapsulated with such films), the effective Fabry-Perot etalon formed by the layer can dominate the thermoreflectance response. This effect has been attributed either to size modulation of the encapsulated material[126] or simply to the change in the reflectance coefficient at the film/substrate[127]. In the former case, the results can be startling, with strong thermoreflectance signals developing even below the bandgap of the imaged material. This effect can also sometimes be used to the advantage of the experimenter. Tessier et al.[128] have shown that a Si_3N_4 encapsulation layer is opaque in the deep UV (λ=240 nm) but has a measurable thermoreflectance coefficient ($\kappa_{TR} \sim 3 \times 10^{-5}$ 1/K). Taking advantage

of this, the entire thermal image can be reconstructed from measurements using only one value of the thermoreflectance coefficient. Additionally, assuming a suitable UV objective is found, the potential exists for imaging the temperature on very small spatial scales (~100nm), since diffraction limit in the deep UV would be correspondingly smaller.

Depending on the specific measurement, a possible cause for concern is the dependence of reflectance on other modulation induced changes in the sample other than the temperature. For example, in semiconductor lasers and amplifiers, the carrier densities in the material can change by a large amount when the device bias is changed. This can change the index of refraction through free carrier absorption, bandgap shrinkage (due to carrier screening), and band-filling effects Figure 87 [129]. The resulting change in index of refraction due solely to the carrier density change will mix with that due to the temperature change, potentially complicating interpretation of the results.

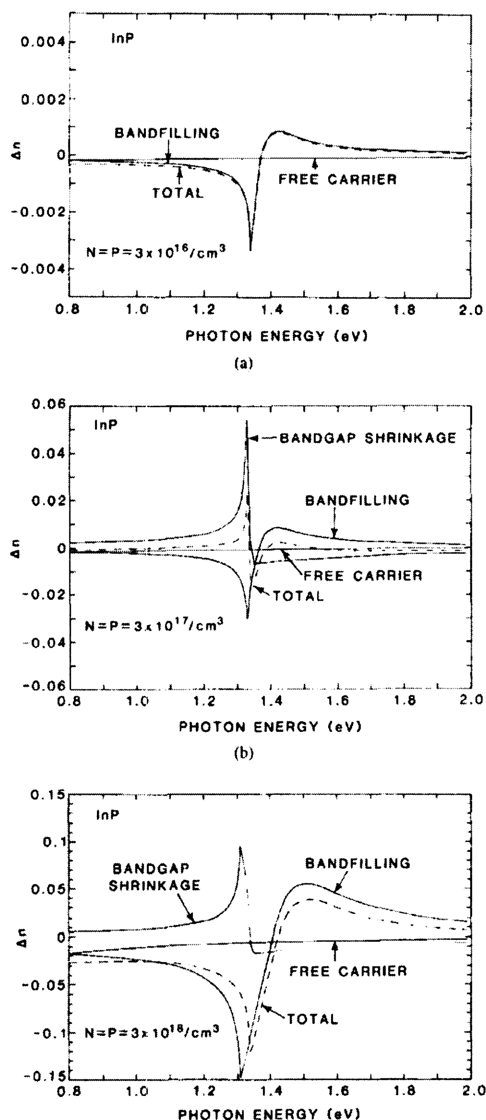


Figure 87. Theoretical calculations of the changes in the index of refraction expected in InP for three doping levels relative to an undoped sample. Three contributing factors (free carrier, band filling, and bandgap shrinkage) cause the changes. (Bennett et al.,1990)

For the measurements in this thesis, we do not expect that carrier densities will change enough to affect the results, owing to the heavily doped or semi-metallic properties of the samples. The light from the diffuse probe LED used to illuminate the sample is not expected to generate sufficient carriers to perturb the thermoreflectance spectrum (this is another advantage of not using intense laser illumination), although we did not experimentally confirm this. Also, as long

as the calibration is performed at the same (steady) illumination as the measurement, this effect would not be expected to affect the temperature measurement.

5.3 Thermoreflectance imaging technique

In this section we review the conceptual trajectory of thermoreflectance imaging in recent years.

The first thermoreflectance imaging systems measured a spatial an image of the thermoreflectance. By using a laser rather than a grating monochromator, the incident beam can be focused on the sample down to near the laser spot-size. If the beam is then rastered across the sample using a piezoelectric stage or movable mirrors, an image can be constructed from the magnitude of the lock-in signal[123,130,131]. Of particular relevance for this thesis, Dilhaire et al. recently used this technique to image the metal top contact of a SiGe/Si superlattice thermoelectric micro-cooler[132], with temperature resolution of below 100 mK.

From our perspective, there are two drawbacks to this approach. The use of coherent light in the measurement can give rise to artifacts (fringes and speckle) in the image if the sample is slightly rough or tilted. Also, the linear scan speed necessary to obtain good thermal resolution can be very slow, on the order of one millimeter per hour[25,133].

An idea for improving the scanning speed of the laser scanning thermoreflectance microscope is to fix the excitation in place, and image the sample with a fixed multi-channel detector array like a CCD camera. The main difficulties with this technique with regards to thermoreflectance are the sensitivity and the frame rate. A type of lock-in measurement sometimes referred to as the "four-bucket" technique has been developed for use with these arrays[133,134]. The basic idea is to trigger the camera to take a picture four times over the course of some phase-locked external modulation (in our case thermal) to the sample. Once data acquisition has terminated, the magnitude and phase of the image's modulation can then be reconstructed from the four distinct phase bins using an analysis similar to that of the four-point discrete Fourier transform. Aside from making the best of the CCD array's slow readout speed, the technique lends itself easily to lock-in imaging by simply accumulating several (identically) phased images into each bin and averaging the results. This is also an economical way to obtain the modulated image, since only 4 images (one for each phase bin) need be kept in memory throughout the

procedure, even if many averages are desired. The basic idea as it applies to thermoreflectance is shown in Figure 88.

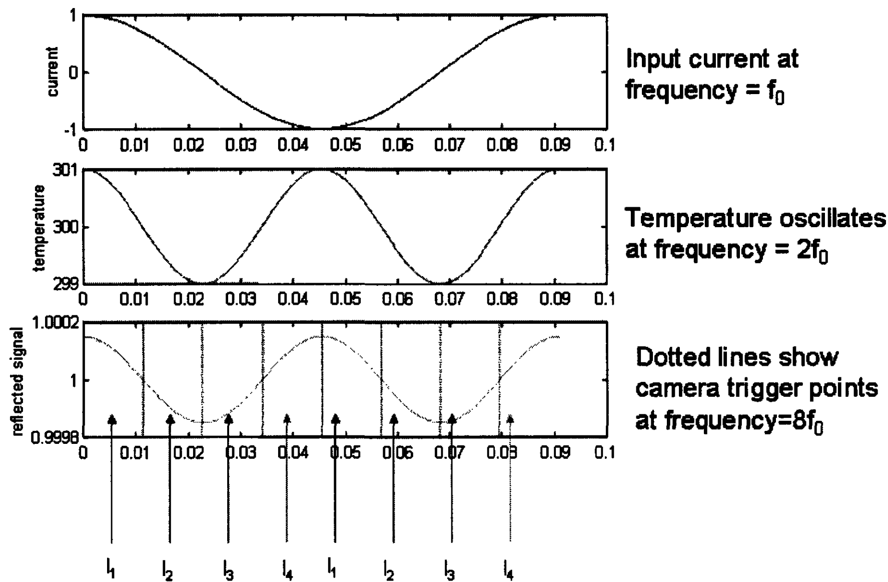


Figure 88. Timing diagram for a measurement of Ohmic (I^2R) heating using the four-bucket technique. The frequency of the heating occurs at 2x the frequency of the incident current, and the camera is triggered at 4x the frequency of the temperature oscillation. The four resulting images are accumulated into bins labeled here as I_1, I_2, I_3, and I_4 for post-processing.

Even using the four-bucket method to lock-in the desired signal, the resolution of the camera may be a cause for concern, since it is nominally set by the camera bit-depth (the number of discrete quantization levels in the analog-to-digital converter).

In an attempt to design a thermoreflectance imaging system with the benefits of both multi-channel imaging and high-resolution, a 16x16 photodiode array was employed by Christofferson et al.[135]. The array was custom built, and relied on 24-bit sigma-delta A-to-D conversion device to match the quantization bin size to the desired thermal resolution (50 mK). A commercial CCD camera was not used because it was believed that the resulting level of quantization (8-14 bits, depending on the model) would be insufficient for recovering thermal signals smaller than around 10 K in magnitude. The array was triggered at a frame rate of 5 kHz. To improve the area imaged, the entire array was placed on a translation stage and stepped around in the image plane for a total scan time of around 6 hours. The resulting set of

images was then digitally stitched together. This camera was used to image the top contacts of superlattice thermoelectric mesas and other devices[136,137].

Despite the apparent limitation presented by the low bit-depth quantization of commercial CCD arrays, they have been used elsewhere with considerable success for thermal imaging. Grauby, Forget, Hole, and Fournier have pioneered much of the work on CCD thermoreflectance microscopy[138]. Their work outlined in this paper forms much of the experimental basis for the measurements of this thesis. A schematic diagram of the thermoreflectance imaging setup used in this thesis is shown in Figure 89.

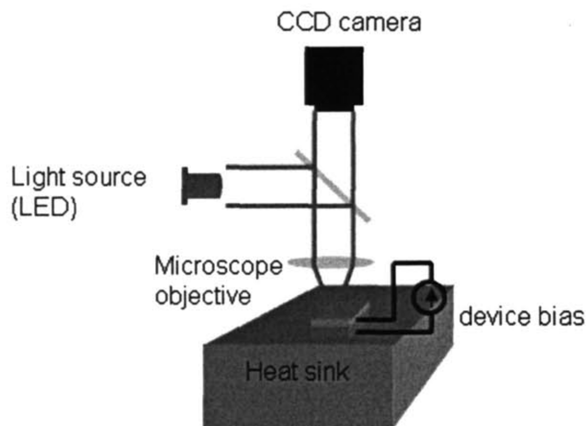


Figure 89. The basic experimental setup used for the thermoreflectance studies of this work.

An LED is used as a monochromatic but incoherent source of light for the measurement. The incident light is coupled into the illumination port of a Mitutoyo microscope using a light pipe, and focused through a long-working-distance objective onto the sample. The reflected light is collected through the objective and imaged onto an Opteon 12-bit CCD camera mounted to the microscope. The temperature of the device is modulated using an external current or voltage drive. The entire measurement setup is secured on a floating optical table, and there are no moving parts or translation stages, minimizing the long term spatial drift inherent in any high spatial resolution microscope.

Associated with the CCD camera is another important contribution to thermoreflectance imaging, and to the work of this thesis in particular. Dietrich Lüerssen, while a post-doc in our group at MIT and at Mt. Holyoke, has developed an excellent thermoreflectance software tool

that we and several other workers in the field have extensively relied on. All of the measurements presented in this thesis were made using his software. The software (TRV2.vi) allows the user to automate the four-bucket lock-in thermoreflectance measurement in the user-friendly Labview programming environment, hiding the customized code needed to control the camera at the firmware level. The was challenging since both read speed and memory management are serious issues for this application; with 493x650 pixels, 12-bit resolution per pixel and an 80 Hz frame rate, the needed data transfer rates from the camera to the controlling computer are in the 10 gigabyte/second range.

Omitted from our measurement apparatus is an innovation used by Grauby et al. designed to broaden the range of thermal phenomena that can be examined using a CCD detector. Aside from its resolution, another serious limitation of a CCD camera in thermal diffusivity measurements or other high speed thermal phenomena is the frame rate of the CCD camera, which is still quite slow (~ 100 Hz) relative to what can be attained by locking in on a single high-speed photo-diode. Grauby et al. proposed a solution to this by heterodyning the measurement. If the sample is modulated at high speed (e.g. 1 MHz) and the illumination is chopped or modulated at a phase-locked frequency just slightly (e.g. 10 Hz) removed from that of the signal, then strobe-like illumination samples the heating in the sample at the right phase intervals so that the desired signal within the measurement band of the camera (triggered at 40 Hz). However, most of the signals of interest to us in this thesis will be inherently slow due to long thermal time constants, or can be made slow without sacrificing the accuracy of the measurement. We have omitted the heterodyning for now, using the simplified setup shown in Figure 89. A slow measurement also allows the camera more time to gather reflected light from the sample, a very important consideration for high resolution (sub 1K) measurements.

Another point made by Grauby et al. is that the measurement can be used to separate the Joule and Peltier components of the heating in a microelectronic structure. In their paper they imaged several polysilicon resistors excited with a harmonic voltage signal at a frequency ω. This resulted in a sinusoidal temperature signal at a frequency 2ω due to Joule heating (I^2R). Using the four-bucket technique, they first triggered their camera at 8ω (four times the frequency of the Ohmic heating signal) and obtained thermal images of the resulting heating, both magnitude and phase. The parabolic dependence of the resulting thermal signal on excitation voltage nicely confirms the existence of Joule heating. The whole measurement is then repeated with a camera sampling frequency of 4ω. At this frequency, the contribution of Joule heating in the

resistor to the measured signal is zero, and the measurement is instead sensitive to any thermal signal that is at the same frequency of the excitation. One such signal is the Peltier effect at the metal/semiconductor contacts on the ends of the resistor, due to the mismatch in the Seebeck coefficient of the polysilicon (from measurements elsewhere[139], α_{poly}~-110 μV/K) and the aluminum (α_{Al}~-2 μV/K). One expects that depending on the direction of the current, there will be alternate heating and cooling at each contact via to the polysilicon resistor, one side cooling as the other heats. Such a signal is observed, and the linearity of the temperature change with drive voltage was confirmed, allowing Grauby et al. to conclude that the Peltier effect had been observed. We are slightly concerned that this measurement may not actually be the Peltier effect but rather the result of non-linear heating in the contact due to the presence of imperfect Ohmic contacts and a voltage source (rather than current source) bias, since a simple back of the envelope calculation predicts that the observed Peltier signal at the contacts will be on the order of 1% to 2% of that of the Joule heating in the resistor. But since their measurement is plotted in arbitrary units rather than in Kelvin or ΔR/R, it is hard to say from their data. At any rate, the important point is that if there is a Peltier signal present, by adjusting the trigger frequency of the camera by a factor of 2, the signal can be picked out separately from the Joule heating. This is obviously very useful in the measurement of the performance of thermoelectric materials. In fact, recent results using a similar setup measuring the surface displacement of the thermoelectric couple (rather than thermoreflectance) have confirmed that this separation is possible in Bi_2Te_3 material[140].

Also of concern during the thermoreflectance measurement is the thermal motion of the sample. There is an entire field of measurement (photothermal deflection microscopy) based on the deflection of an incident laser beam from a surface distorted by a thermal disturbance. From the standpoint of thermoreflectance imaging, such sample distortion is generally regarded as spurious noise, since the physics of the resulting signal has nothing to do with the band-structure/broadening model discussed in Section 5.2. Other researchers have taken pains to monitor both the mechanical and the thermoreflectance signal separately[138,141], for example to obtain the temperature and displacement of a MEMS heater. For a number of reasons, we do not do that. First, we use incoherent sources (LEDs). These do not suffer from the speckle or fringe type effects common to laser light under most circumstances. Second, we measure all devices at normal incidence, such that thermal expansion or contraction of the device under study does not cause much deflection of the light. For higher numerical apertures this is more of a concern, as has been pointed out by Dilhaire et al.[142], who find that with an NA=0.8

microscope objective the image intensity (and hence the measured thermoreflectance) of a sample that moves up or down 300 nm relative to the focal plane of the microscope will change by 5×10^{-4} due to the changing amount of collected light. The error vs. displacement is parabolic in shape, so that very small displacements cause almost no error. In any case, the importance of any sample motion can be experimentally determined by moving the objective of the microscope a small distance vertically and examining and change in the dc reflectance of light from the sample. The error is related to the "depth of focus" specification of the microscope objective. The depths of focus for the Mitutoyo "M Plan APO Long Working Distance" objectives used in this thesis were 0.6 μm for the 100x, 0.9 μm for the 50x, and 1.6 μm for the 20x. Finally, whether the origin of the signal is due to thermoreflectance or another thermal effect, the intent of the calibration procedures developed is to obviate the need to know the exact physics behind the reflectivity change of the sample. The calibration is important in any case, since the doping, orientation, and effects of any sample specific surface-oxides are not necessarily the same as those in the literature values of the thermoreflectance coefficient of a given material.

Although quantitative thermal images have been produced from the thermoreflectance technique, an important question is unanswered: what are the ultimate attainable spatial and temperature resolutions? Both issues were experimentally investigated by Lüerssen et al., and a method for determining the validity of a measurement over a given region using the measured phase noise was provided[143]. Lüerssen argues that the so called Sparrow criterion is the proper criteria for evaluating the resolution of an imaging system, and verifies his 225 nm optical resolution with an optical measurement of a small feature using a NA= 0.8 microscope objective and a blue (467 nm) illumination wavelength. There are two possible issues with this conclusion.

First, an optical image was used to confirm the resolution rather than a thermal one, since submicron thermal test structures were not readily available. Second, under a long measurement, there may be thermal drift of the microscope[144]. This drift motion likely has a power-law distribution from which the error does not average out over time. While such noise in the temperature of the sample is mostly rejected using the four-bucket lock-in technique, this lateral drift simply causes blurring of the images, and over long enough periods can limit the resolution of an imaging system. However, since the importance of this noise is a sensitive function of the details of a given thermoreflectance setup, it is something probably best measured by individual practitioners of the technique by simply performing long acquisitions on

fine features and analyzing the resulting image blurring. Also, if the drift is a problem, in principle the images can be corrected in software down to close to the optical resolution[145], although thus far, this has not been necessary for our setup. As a final note, even the Sparrow criterion is not the ultimate limit for the resolution of thermoreflectance imaging. Recent work using scanning near-field optical microscopy, although still in its initial stages, offers the prospect of imaging below the diffraction limit[130,146]. A lateral resolution of 100nm has been claimed, although quantitative temperature measurements have not yet been performed.

Lüerssen et al. also provide empirical evidence for 10 mK temperature resolution on a gold surface, using an illumination wavelength optimized for the material. Experimental confirmation of such high temperature resolution is very important; while the exact magnitude of the measured temperature is not clear from their paper discussed above[138], other results of Grauby et al. seem to focus on large temperature differences in the 10s of K, with substantially less resolution than Lüerssen[147]. This high temperature resolution also contradicts the view that the resolution of CCD-based thermoreflectance imaging should be limited to the quantization bin size of the camera[135]. In fact, the equivalent bit depth attained by Lüerssen et al. was actually close to 18 bits rather than the 12 bits offered nominally by the camera. Lüerssen et al. provide a qualitative explanation for this based on the idea of stochastic resonance, an idea with broad applications in nonlinear physical systems in general[148]. Put more simply, the extra resolution results from the repeated averaging of the quantized signal, in the presence of noise sufficiently large so as to knock the camera pixel's signal around between adjacent quantization levels. How this works is shown in Figure 90.

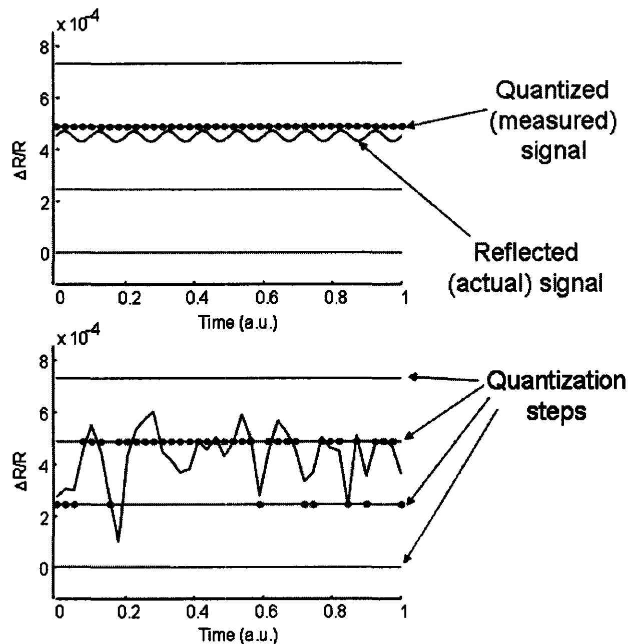
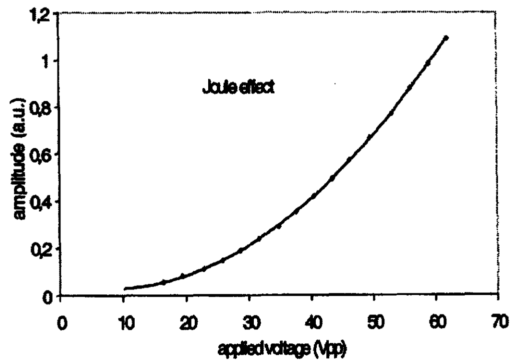
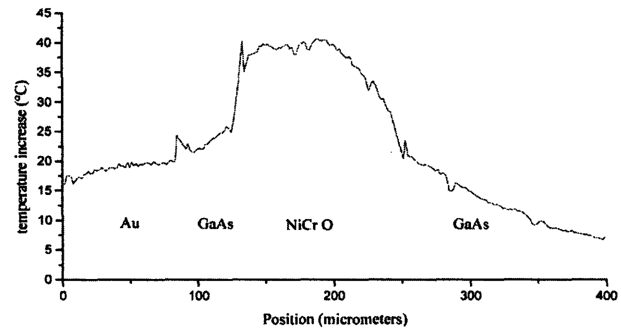


Figure 90. Measurement below the quantization threshold using oversampling and averaging. With no noise (top), the dynamic range of the measurement is limited by the quantization step size, regardless of any averaging. With sufficient noise (bottom), the dynamic range can be improved through averaging, allowing a signal between the quantization levels to be measured.

It is amusing that any measurement should be improved by the presence of noise. Without boosting the resolution of the camera using this averaging, the thermal resolution for a typical thermoreflectance measurement would be at most $\frac{\kappa_{TR}^{-1}}{2^b}$, where κ_{TR} is the thermoreflectance coefficient and b is the bit-depth of the camera. For a typical material κ_{TR} =1 x 10^{-4} 1/K and for a typical camera b=12, limiting the resolution to around 2.5 K. The best results in the literature that we are aware of for CCD thermoreflectance imaging from near the start of our research program, demonstrated the potential of the technique but left the question of temperature resolution and accuracy open. Figure 91 depicts two of results from Grauby et al. The left most image plots the parabolic heating profile in a chip resistor, but uses arbitrary units. The right-most image depicts the heating in a different resistor has a temperature axis, but the accuracy and resolution (not explicitly stated in the paper) seem to be on the order of a few degrees K.



from "High resolution photothermal imaging of high frequency phenomena using a visible charge coupled device camara associated with a multichannel lock-in scheme"
S. Grauby, B. Forget, S. Hole, D. Fournier.
Review of Scientific Instruments. v70, (1999)



from "Quantitative thermal imaging with CCD array coupled to an heterodyne miltichannel lock-in detection"
S. Grauby, G. Tessier, V. Rachet, S. Hole, D. Fournier, Analytical Sciences, v17 (2001)

Figure 91. Results using CCD thermoreflectance imaging from Grauby et al. (1999, 2001)

Despite the empirical sub-quantization level measurements of Lüerssen and our qualitative understanding of how it is obtained, there were (at the start of this work) no published results describing the theory of how the resolution enhancement is obtained. The most important questions for an experimenter are how long one must measure to obtain a certain temperature resolution, what are the resulting error bars on the measurement, and how much noise is present in the camera (since according to the stochastic resonance theory there ought to be an optimal noise level for signal recovery). Developing this theory is the goal of the next section.

Before moving on, it is important to note that there exist several other methods for obtaining thermal images of a sample. The most familiar is that of infrared imaging. An infrared image is constructed by scanning (or photographing) a sample with an infrared photodetector (or IR camera). Possible detectors and their range of wavelength sensitivity include PtSi Schottky (1.1-5 μm), InSb (3-5 μm), HgCdTe (3-5 μm or 8-10 μm), and GaAs/AlGaAs quantum well devices (6-20 μm). Provided the emissivity of the material is known (often obtained using a calibration procedure similar to that used for thermoreflectance), monitoring the blackbody emissivity of a sample with a cooled detector allows the temperature to be measured with excellent thermal resolution (~ 10 mK). The frame rates of good IR cameras are similar to those of CCD arrays (~50 Hz). Unfortunately, the practical spatial resolution is ultimately limited to around 4 μm by the large IR detection wavelength, the low spectral emissivity of objects near room temperature, and the physical limitations on the IR camera objective and detector pitch inside the camera. This is too large for imaging a thermoelectric thin-film in the transverse

direction, and so this direction was not pursued for this thesis. We note than an additional attraction of standard CCD arrays (sensitive in the visible) over state-of-the-art liquid nitrogen or Stirling-cooled focal plane arrays (sensitive in the infrared) is around a two order of magnitude reduction in price.

Another novel technique designed for sub-micron back-plane imaging is the SILS (solid immersion lens) technique. The basic idea is that a solid lens is placed on the sample (or fabricated there), allowing the effectively numerical aperture of the image onto an infrared camera. Resolutions of 1.4 μm have been experimentally confirmed, with the theoretically possible resolution in the submicron range[149]. For our purposes, the measurement is less attractive, due to the need to construct a special structure in optical (and therefore) thermal contact with the device under test, possibly altering the temperatures and heat thereof.

There have been numerous special techniques developed specifically for thermal metrology in the nanoscale. Cahill, Goodson, and Majumdar provide an excellent review of these techniques[150]. Arguably the most interesting technique discussed there for the type of measurements pursued in this thesis is scanning thermal microscopy (SThM)[151], which has demonstrated a low spatial resolution of < 50nm. The basic idea is to scan the surface of a material using an atomic force microscope (AFM) with a special thermocouple tip. Although the spatial resolution of the technique is impressive, a possible issue for this thesis (aside from the construction of the AFM and tip fabrication) is calibration. The measurement is very sensitive to the details of the tip-sample interaction, and obtaining trustworthy measurements of absolute temperature in the sub-1K resolution range has been difficult.

5.4 Theory of sub-quantization thermoreflectance imaging

There are several important reasons to carefully work through the thermoreflectance measurement theory. A good understanding of the noise floor is necessary to trust the results of any measurement, and to obtain sensible error bars. Specifically for a lock-in measurement, a key question is how much averaging is required before the results can be trusted. For a typical lock-in amplifier with a time constant of a few seconds, trial-and-error can suffice, but for lock-in imaging techniques (when scanning or camera read rates can limit measurement speed to hours) this method is prohibitively time consuming. Also, for measurements using an A-to-D converter, a key theoretical question is what effect (if any) the dynamic range limit set by the

minimum quantization bin size has on the measurement. In this section, a theory of the thermoreflectance measurement is developed and these questions are answered. The theory is then tested against measurements and found to be in satisfactory agreement.

The 4-bucket technique for thermoreflectance measurements accumulates CCD counts into 4 distinct bins, one for each phase of the temperature modulation. The counts in each bin for an individual pixel (indexed by x and y) can be described by

$$I_k(x,y) \equiv \sum_{i=1}^{N} \left\lfloor \left(\frac{4}{T} \int_{(4i+k)\frac{T}{4}}^{(4i+k+1)\frac{T}{4}} [c(x,y) + \Delta(x,y)\cos(\omega t + \phi(x,y) + \psi)] dt \right) + d(x,y) \right\rfloor \quad (5.4.1)$$
$$k \in \{1,2,3,4\}$$

The total thermoreflectance signal is quantized (square L-brackets) and then accumulated into each of the four bins for N iterations of the thermal cycle. Here the term c represents the steady (dc) part of the reflected light incident on the pixel, and the term Δ represents the thermoreflectance signal, assumed to be sinusoidal at a frequency ω. There is also some noise present prior to quantization, described by d. The phase of the harmonic heating is described by ϕ, and we include an arbitrary uniform phase offset Ψ. To extract the magnitude and phase of the thermoreflectance signal, the following quantities are formed from the buffers once data acquisition is completed:

$$\left|\frac{\Delta R}{R}\right|_{ideal} = \frac{\pi}{\sqrt{2}} \frac{\sqrt{(I_1 - I_3)^2 + (I_2 - I_4)^2}}{I_1 + I_2 + I_3 + I_4} = \frac{\Delta}{c} \quad (5.4.2)$$

$$\phi_{ideal} = \arctan\left(\frac{I_1 - I_2 - I_3 + I_4}{I_1 + I_2 - I_3 - I_4}\right) - \psi = \phi \quad (5.4.3)$$

The last equalities in Eq. 5.4.3 and Eq. 5.4.4 hold if the effects of quantization and noise are ignored, and so in some sense are the "ideal" measurement result. Up to an arbitrary phase offset, they are similar to what has been reported elsewhere[138].

In order to attain thermal images of a sample, many images must be averaged so that the signal to noise ratio of the thermoreflectance magnitude and phase are reduced. One "quick and dirty" way to tell whether enough averaging has been performed is to examine the phase image of a finished measurement. Figure 92 shows the results of a thermoreflectance measurement on a polysilicon resistive line in a silicon substrate. The polysilicon in the image is one uniform color, indicating that the phase is a constant there, and is well-defined. The surrounding silicon and

the aluminum lines which cross the polysilicon have smaller thermoreflectance signals that have not yet converged, and so the phase appears speckled.

Figure 92. A thermoreflectance phase image illustrating partical phase convergence. The solid area depicts a polysilicon resistive snake in a Si substrate. The speckled signal elsewhere is a result of random phase due to insufficient signal power.

To examine the question of how many iterations are required to achieve a given temperature resolution, and the specific form of the convergence of the measurement, we investigate the error in the measurement by treating $\left|\frac{\Delta R}{R}\right|$ and ϕ as random variables with probability density functions (pdfs) that are seeded by the input noise in each pixel (the random variable d). The pixel noise is a combination of shot noise and thermal read-out noise[152], both zero mean, white Gaussian noise processes, so that

$$p_d = \mathcal{N}\left(0, \sigma_d^{\,2}\right), \qquad (5.4.5)$$

where the notation $\mathcal{N}\left(\mu, \sigma^2\right)$ denotes a Gaussian pdf with mean μ and variance σ. The noise σ_d is assumed to be expressed in units of CCD counts, although we allow fractional values of σ_d because the noise here is being modeled before the quantization step in the CCD camera. In terms of the number of CCD counts in the signal, the standard deviation of the noise can be written as

$$\sigma_d = \alpha\sqrt{c} + \beta . \qquad (5.4.6)$$

Here α is picked to model the shot noise of the incoming photons and dark current, and β models all of the noise which does not depend on the incoming signal intensity (e.g. thermal noise).

During the measurement, the light signal and noise are then quantized by the camera wells. This is a nonlinear process whose output on the measurement generally depends nontrivially on the signal[153]. However, if the noise source in the incoming signal is larger than the quantization bin size of the camera ($\sigma_d > 1$), the effects of quantization noise can be represented as an additive (uncorrelated) white noise process of standard deviation $\sigma_{quant} = \frac{1}{\sqrt{12}}$ (see Appendix 1). With regards to its effect on the output, the quantization noise is then exactly like the thermal noise, and can be absorbed into the β coefficient in the expression for the pixel noise (Eq. 5.4.6).

Each of the random variables I_k in Eq. 5.4.2 are constructed from the sum of N successive exposures of the CCD camera, and can hence be described by Gaussian pdfs according to

$$p_{I_k} = \mathcal{N}\left(\mu_k, N\sigma_d^{\,2}\right). \qquad (5.4.7)$$

Because the noise in each exposure is uncorrelated with the others, the noise variances for each exposure add, resulting in the factor of N. The means μ_k are easily calculated:

$$\mu_k = N\left(c - \frac{4\Delta}{\pi\sqrt{2}}\left(\cos\left(\frac{\pi k}{2} - \frac{\pi}{4}\right)\sin(\phi+\psi) + \sin\left(\frac{\pi k}{2} - \frac{\pi}{4}\right)\cos(\phi+\psi)\right)\right) \qquad (5.4.8)$$

For the (more realistic case) when the CCD camera acquires data for each image for less than the full quarter-period of the temperature oscillation, the expression in Eq. 5.4.8 is a little more complicated, but is easily calculated. In any case, the difference between the full quarter-period integration of Eq. 5.4.8 and the opposite extreme (delta function sampling) is only around 10%.

To calculate the probability distribution of the observed output magnitude and phase, the Gaussian distribution of Eq. 5.4.6 is propagated through Eq. 5.4.3 and Eq. 5.4.4. For thermoreflectance measurements the constant background reflectance signal c is far larger than both the noise and the modulated thermoreflectance signal, $\mu_k \approx Nc >> \sqrt{N\sigma_d^{\,2}}$, and the variance of I_k in the denominator of Eq. 5.4.3 can be ignored. In the numerator, because the differences between I_k are taken, this is not the case. Eq. 5.4.3 can be rewritten

$$\left|\frac{\Delta R}{R}\right| = \sqrt{\left[\frac{\pi}{\sqrt{2}}\frac{I_1 - I_3}{\sum_k \mu_k}\right]^2 + \left[\frac{\pi}{\sqrt{2}}\frac{I_2 - I_4}{\sum_k \mu_k}\right]^2} = \sqrt{\left[\frac{\pi}{\sqrt{2}}\frac{I_1 - I_3}{4Nc}\right]^2 + \left[\frac{\pi}{\sqrt{2}}\frac{I_2 - I_4}{4Nc}\right]^2}. \qquad (5.4.9)$$

Since the I_k are statistically independent random variables, the bracketed terms are themselves Gaussian random variables, allowing new independent random variables to be defined:

$$A = \frac{\pi}{\sqrt{2}}\frac{I_1 - I_3}{4Nc}, \quad p_A = \mathcal{N}(\mu_A, \sigma^2), \qquad (5.4.10)$$

$$B = \frac{\pi}{\sqrt{2}}\frac{I_2 - I_4}{4Nc}, \quad p_B = \mathcal{N}(\mu_B, \sigma^2), \qquad (5.4.11)$$

where

$$\mu_A = \frac{\pi}{\sqrt{2}}\frac{\mu_1 - \mu_3}{4Nc}, \qquad (5.4.12)$$

$$\mu_B = \frac{\pi}{\sqrt{2}}\frac{\mu_2 - \mu_4}{4Nc}, \qquad (5.4.13)$$

$$\sigma^2 = \frac{\pi^2}{2}\frac{2N\sigma_d^{\,2}}{(4Nc)^2} = \frac{\pi^2}{16}\frac{\sigma_d^{\,2}}{Nc^2}. \qquad (5.4.14)$$

The thermoreflectance is then just

$$\left|\frac{\Delta R}{R}\right| = \sqrt{A^2 + B^2}. \qquad (5.4.15)$$

Equation 15 has a clear geometrical interpretation: since A and B are regarded independent (orthogonal) random vectors, then the thermoreflectance is their magnitude. This suggests a method for finding the pdf (probability density function) of the thermoreflectance.

The joint Gaussian pdf of A and B is:

$$p_{A,B}(A,B) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(A-\mu_A)^2 + (B-\mu_B)^2}{2\sigma^2}\right). \qquad (5.4.16)$$

By changing variables according to $A = r\cos\theta$ and $B = r\sin\theta$, one finds the joint pdf in terms of the new variables:

$$p_{r,\theta}(r,\theta) = \frac{r}{2\pi\sigma^2}\exp\left(-\frac{1}{2\sigma^2}\left[(r\cos\theta - \mu_A)^2 + (r\sin\theta - \mu_B)^2\right]\right). \qquad (5.4.17)$$

By integrating this over all θ one finds the pdf for r (the thermoreflectance), given by:

$$p_r(r) = \frac{r}{\sigma^2}\exp\left(-\frac{r^2+\mu_A{}^2+\mu_B{}^2}{2\sigma^2}\right)I_0\left(r\frac{\sqrt{\mu_A{}^2+\mu_B{}^2}}{\sigma^2}\right), \tag{5.4.18}$$

where I_0 denotes the 0th order modified Bessel function. This is the so-called Rician distribution[154]. It is reassuring to note that in the limit of low noise (small σ), the Rician collapses to a delta function distribution centered on the value

$$\left|\frac{\Delta R}{R}\right|_{ideal} = \sqrt{\mu_A{}^2+\mu_B{}^2} = \frac{\Delta}{c}, \tag{5.4.19}$$

as stated in Eq. 5.4.2. The pdf of Eq. 5.4.17 can also be integrated to find a closed expression for the phase noise. This is done in Appendix 2.

The pdf of Eq. 5.4.18 completely describes the statistics of the measured thermoreflectance magnitude, and allows for the theoretical determination of the number of iterations required to accurately measure a given temperature difference, as long as the mean camera counts (c) and the input noise characteristics (Eq. 5.4.6) are known. The moments are given by:

$$E\left[\frac{\Delta R}{R}\right] = \sqrt{2}\sigma\exp\left(-\frac{\mu_A{}^2+\mu_B{}^2}{2\sigma^2}\right)\left[\frac{1}{2}\sqrt{\pi}\left(1+\frac{1}{2}\frac{\mu_A{}^2+\mu_B{}^2}{\sigma^2}\right)\exp\left(\frac{\mu_A{}^2+\mu_B{}^2}{4\sigma^2}\right)I_0\left(\frac{\mu_A{}^2+\mu_B{}^2}{4\sigma^2}\right)+ \right. \\ \left. \frac{\sqrt{\pi}\left(\mu_A{}^2+\mu_B{}^2\right)}{4\sigma^2}\exp\left(\frac{\mu_A{}^2+\mu_B{}^2}{4\sigma^2}\right)I_1\left(\frac{\mu_A{}^2+\mu_B{}^2}{4\sigma^2}\right)\right] \tag{5.4.20}$$

$$E\left[\left(\frac{\Delta R}{R}\right)^2\right] = 2\sigma^2+\mu_A{}^2+\mu_B{}^2. \tag{5.4.21}$$

ectance measurement for varying nu plots an example of the behavior of a measurement as the number of measured iterations (N) is increased.

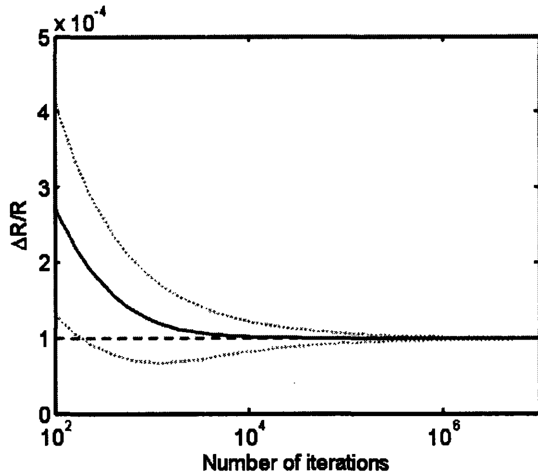


Figure 93. A thermoreflectance measurement for varying numbers of iterations. The dashed line is the ideal thermoreflectance whose measurement is desired. The solid line describes the expected value (first moment) of the thermoreflectance. The dotted lines describe the +/- 1 sigma error bars of the measurement.

Of interest to the experimenter is the error in the thermoreflectance measurement, which can be well approximated by the sum of two errors ε_{offset} and ε_{random}. ε_{offset} is the difference of first moment of the Rician pdf in Eq. 5.4.18 from the ideal signal in Eq. 5.4.19. ε_{random} is the standard deviation of the distribution, and is related to the random error of the measurement. The offset error is defined by

$$\varepsilon_{offset} \equiv E\left[\frac{\Delta R}{R}\right] - \left.\frac{\Delta R}{R}\right|_{ideal} = E\left[\frac{\Delta R}{R}\right] - \sqrt{\mu_A^{\ 2} + \mu_B^{\ 2}}. \qquad (5.4.22)$$

The random error is given by

$$\varepsilon_{random} \equiv \sqrt{E\left[\left(\frac{\Delta R}{R}\right)^2\right] - E\left[\frac{\Delta R}{R}\right]^2}. \qquad (5.4.23)$$

If there is no thermoreflectance signal present ($\mu_A^2+\mu_B^2=0$), the expression for ε_{random} can put in terms of the fundamental noise σ_d of the CCD pixel:

$$\left.\varepsilon_{random}\right|_{\mu_A^2+\mu_B^2=0} = \left(\sqrt{2-\frac{\pi}{2}}\right)\sigma = \left(\sqrt{2-\frac{\pi}{2}}\right)\frac{\pi}{4}\frac{\sigma_d}{c\sqrt{N}} = \left(\sqrt{2-\frac{\pi}{2}}\right)\frac{\pi}{4}\frac{\left(\alpha\sqrt{c}+\beta\right)}{c\sqrt{N}}. \qquad (5.4.24)$$

By measuring $\frac{\Delta R}{R}$ with no thermoreflectance signal for a variety of count levels c and iterations N, one can measure $\left.\varepsilon_{random}\right|_{\mu_A^2+\mu_B^2=0}$, and hence determine the fundamental input noise constants α and β. Once these are known, the accuracy of thermoreflectance measurements

with a specific CCD camera can be analytically predicted as a function of c and N. Alternatively, one can make a direct measurement of the camera noise parameters α and β by measuring of the pixel noise in normal images as a function of N and c [152].

The errors are plotted in Figure 94 below as a function of c and N. For a given level of dc camera signal c, this chart can be used to determine the number of iterations N required to achieve a given accuracy. The color bar to the right of the figure corresponds to the log base 10 of the error magnitude.

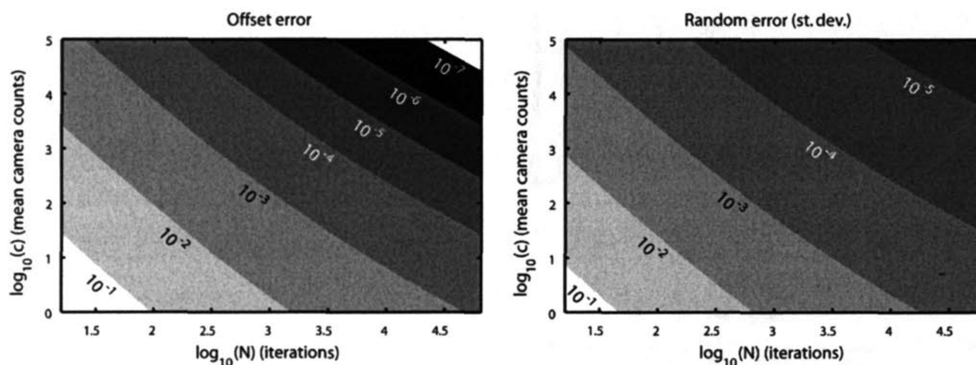


Figure 94. Example of plots of ε_{offset} and ε_{error} as a function of iterations and CCD counts. The plots here allow the error to be predicted given a value α=0.207 and β=7.650 for the fundamental noise in the camera, and a thermoreflectance signal of magnitude $\Delta R/R=10^{-4}$.

The results of this section have been incorporated into a MATLAB function (see Appendix 6) that returns the offset error and random error in a thermoreflectance measurement given the magnitude of the thermoreflectance signal $\Delta R/R$, the average number of dc camera counts c, the number of measurement iterations N, and the two camera noise parameters.

5.5 Experimental verification of sub-quantization imaging theory

As direct experimental confirmation that meaningful signals smaller than the least significant bit of the detector can be handled, we measured the linear temperature dependence of boron-doped Si resistor with Ohmic heating power (Figure 95). The current into the resistor was modulated sinusoidally at 6 Hz, resulting in a modulated temperature (and reflectance) signal at 12 Hz due to Joule heating. The CCD camera trigger was locked to the reflectance signal at 48 Hz. The doped resistor's thermoreflectance coefficient κ_{TR} =2.20 x 10^{-4} 1/K was measured by

using a micro-thermocouple as a reference temperature monitor. The mean level of camera counts for the measurements was 1550, giving rise to a quantization limit temperature level of 2.93 K. The data presented in Figure 95 clearly show the linear dependence of the measured temperature on the electrical power, even though all signals are measured far below the temperature corresponding to the quantization limit of the camera. For most of these measurements $N=10^5$ iterations were used, resulting in a measurement time of 2.3 hours. Less stringent measurements can be performed in minutes. The standard deviation of the error from the linear fit in Figure 95 is 18mK. In separate measurements performed on gold using 467nm illumination and described in detail elsewhere[143], the thermoreflectance coefficient is as high as $\kappa_{TR} = 3.3 \cdot 10^{-4}\,K^{-1}$, allowing a per-pixel temperature resolution of 10mK for the parameters mentioned above (1.5mK at 900nm spatial resolution).

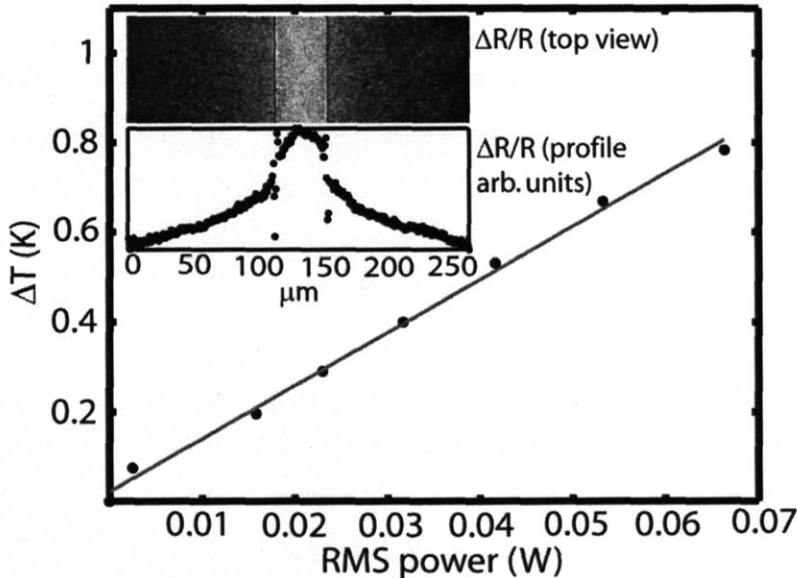


Figure 95. Experimental confirmation of sub-quantization level imaging. Thermal images (top inset) of a boron diffused resistor in Si were used to find the temperature profile (lower inset) across the resistor. The average temperature inside the resistor obeyed the expected linear relationship with electrical power. The temperature corresponding to the quantization limit was 2.93 K; all data points are below that.

It is also desirable to obtain direct experimental verification of the detailed noise theory presented in the previous section. To verify the theory, a 25 pixel (5 x 5) area of uniform temperature excitation was imaged for varying numbers of iterations by Dietrich Lüerssen. This was done for two independent levels of illumination and temperature excitation. Separate measurements and the use of Eq. 5.4.24 determined the values of the noise parameters:

α=0.207 and β=7.650. Figure 96 plots two examples of the behavior of a measurement as the number of measured iterations (N) is increased, one for the larger thermoreflectance signal and one for the smaller signal. The grey dotted lines are the mean measured thermoreflectance of the 25 pixels, the black solid line depicts the theoretically expected thermoreflectance, and the dotted lines denote the theoretical 1-sigma error bars to the measurement. We note that it is only for sufficiently large enough iterations N that the mean of the measurements converge. For insufficient averaging (small N), it is clear from Figure 96 that the (initially zero-mean) Gaussian pixel noise can give rise to a large systematic offset error in the measured thermoreflectance. The understanding and mitigation of this error by proper application of the noise theory is therefore very important for quantitative thermoreflectance imaging at high levels of temperature sensitivity.

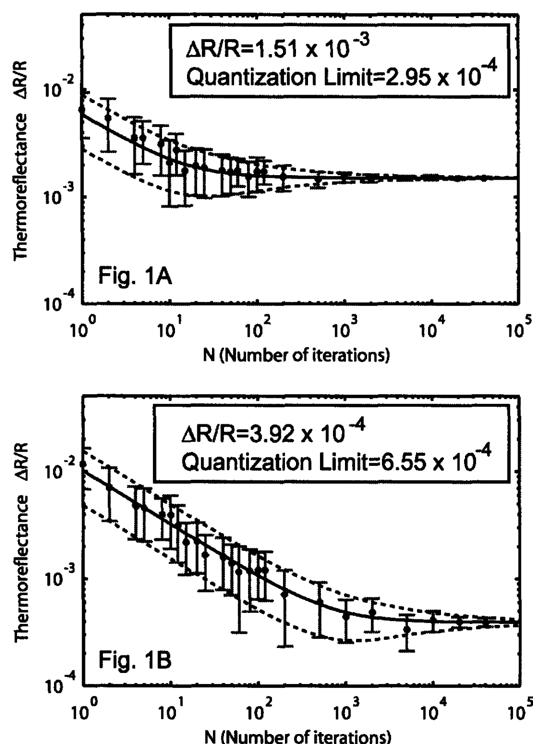


Figure 96. Example of plots of ε_{offset} and ε_{error} as a function of iterations and CCD counts. The plots here allow the error to be predicted given a value α=0.207 and β=7.650 for the fundamental noise in the camera, and a thermoreflectance signal of magnitude $\Delta R/R=10^{-4}$. The measured signal and measured +/- 1 sigma error bars are shown in gray. A black solid line depicts the theoretical signal, and the black dotted lines depict the theoretical +/- 1 sigma error limits.

5.6 Thermoreflectance calibration measurements

Like most quantitative thermal imaging techniques, accurate determination of the temperature of an imaged surface requires a calibration measurement. While published values for some thermoreflectance coefficients are available, variations in samples, doping, surface passivation/oxide, crystal orientation, and the spectral content of the light source often make individual calibration necessary. Calibration is most straightforward on a sample whose thermal mass is large enough to accurately measure using a microthermocouple (25μm diameter), and whose thermal mass is small enough to undergo periodic changes in temperature of a few K at a moderate (0.2-10 Hz) frequency. For most samples of interest, this is a reasonable constraint, and pieces of semiconductor wafers end up being particularly convenient. The idea behind the calibration is shown in Figure 97.

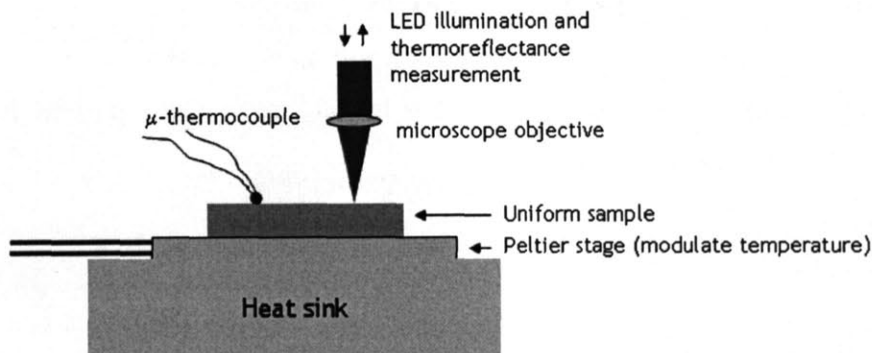


Figure 97. Thermoreflectance calibration procedure used in this work. A thermoreflectance measurement on a uniform area of the sample is performed, while on an adjacent area, a microthermocouple monitors the surface temperature. A Peltier stage is driven with a sinusoidal current to modulate the sample temperature.

The sample's temperature is sinusoidally modulated using a Peltier cooler mounted on a heat sink of large thermal mass. A cooler with short legs and a fast thermal response time is generally most convenient. A thermogram (thermal image) of a uniform area of the sample is obtained while the amplitude ΔT of the surface temperature modulation is simultaneously acquired using a microthermocouple adjacent to the image and stuck to the sample using heat sink grease. The thermoreflectance amplitude is then found by calculating the thermoreflectance $\Delta R/R$ from a pixilated average of I_1, I_2, I_3, and I_4 over the region of interest in

the image. The thermoreflectance coefficient $\kappa_{TR} = \frac{\Delta R/R}{\Delta T}$ is then found. It is very important to first average the I_1, I_2, I_3, and I_4, and then calculate the thermoreflectance from Eq. (5.4.9). If instead, the thermoreflectance $\Delta R/R$ is first calculated for each pixel and then a pixel average of $\Delta R/R$ is computed, there will be a much larger offset error. The reason is that whenever $\Delta R/R$ is calculated, an offset is introduced whose magnitude is related to the underlying noise variance as discussed in the previous section (Eq. 5.4.22). Subsequent averaging of the $\Delta R/R$ will not reduce this error at all. One should therefore use an I_1, I_2, I_3, and I_4, with the lowest possible noise before calculating $\Delta R/R$, which entails averaging them first to reduce their variance.

An example of the measurement of the thermoreflectance coefficients is shown in Figure 98. Using convenient test structures on a Si CMOS-process wafer, the thermoreflectance coefficients of three materials can be extracted using the procedure described above. In addition, the phase of the image can be used to extract the relative signs of the thermoreflectance coefficient (in this case, un-doped Si was known to have a positive thermoreflectance coefficient). The negative coefficient of the polysilicon region shows up as a 180 degree phase shift in the thermoreflectance image.

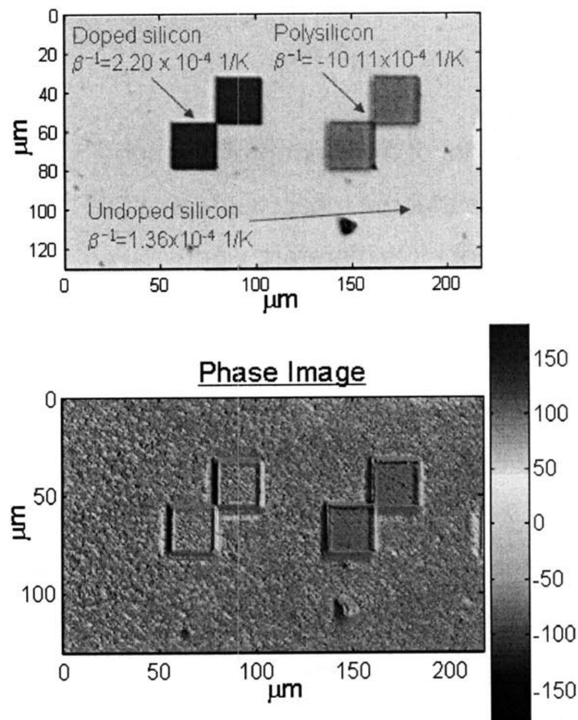


Figure 98. Example of the simultaneous acquisition of three thermoreflectance coefficients (calibrations) using one thermoreflectance image and a microthermocouple (not shown). The top photo shows a test structure on a Silicon wafer and the three areas used. The bottom picture is the phase image of the thermoreflectance measurement.

Also visible in the phase image are some artifacts due to sample roughness and a large dust particle (above the 150 μm mark in the lower right quadrant). If a sample is rough or has other sharp changes in features, any sample motion that is periodic with the thermal excitation can cause artifacts, since the modulation of the illumination of a single pixel due to sample motion is indistinguishable from that due to true thermoreflectance. The effect is instantly recognizable in the phase image though, since edges and points develop a characteristic mirror phase image immediately adjacent to them. In fact, a phase image of a completely rough sample looks like a two-colored speckle, and a plot of the histogram of phases for such a material reveals a bimodal phase distribution. The artifacts in Figure 98 are due to the slight motion (~ μm) of the sample as the relatively large (1 mm thick) Peltier stage beneath heats and cools. They can be alleviated by using a smoother sample, or going to a weaker magnification, but caution must be used when interpreting images near sharp material or geometrical boundaries. Measurements

of devices whose temperatures can be modulated using an electrical bias (as opposed to calibration measurements which are done on a large Peltier stage) are less prone to this artifact because the volume of the sample undergoing heating and cooling is much smaller, resulting in less motion due to thermal expansion and contraction.

The calibration procedure described above for the measurement of the thermoreflectance coefficient is referenced to a micro-thermocouple. This means that the measurement of the temperature is only as good as the thermocouple measurement. It is therefore worth considering the accuracy of the thermocouple measurement.

There are a number of ways in which thermocouples can introduce measurement errors [81,155]. The thermal parasitic load of the thermocouple's wires, the heat transfer from the tip, and the finite heat capacity (charging time) can all act to decrease the measured temperature relative to the true temperature. So too can the existence of a finite thermal resistance between the couple and the measured surface. The use of a microthermocouple is expected to minimize all of these errors. To verify that this is true, we attempted to verify a calibrated thermoreflectance measurement on the doped silicon resistor of Figure 95 using a 3ω measurement of the temperature in the resistor. By performing a 4-point resistivity measurement on the resistor while slowly ramping the resistor's temperature, we found the temperature coefficient of the resistor to be α_{TC}=0.00142 1/K. Using this value, a 3ω measurement was performed using a custom made preamplifier differencing circuit and a commercial lock-in amplifier (see section 4.2 for a discussion of the 3ω measurement). Under an identical bias (10 mA → 56.3 mW) and current frequency (4 Hz), the thermoreflectance image of the thermal modulation at 2ω was also taken. The temperature profile across the resistor was found according to the thermoreflectance coefficient measured using the thermocouple calibration described above. The results are shown in Figure 99.

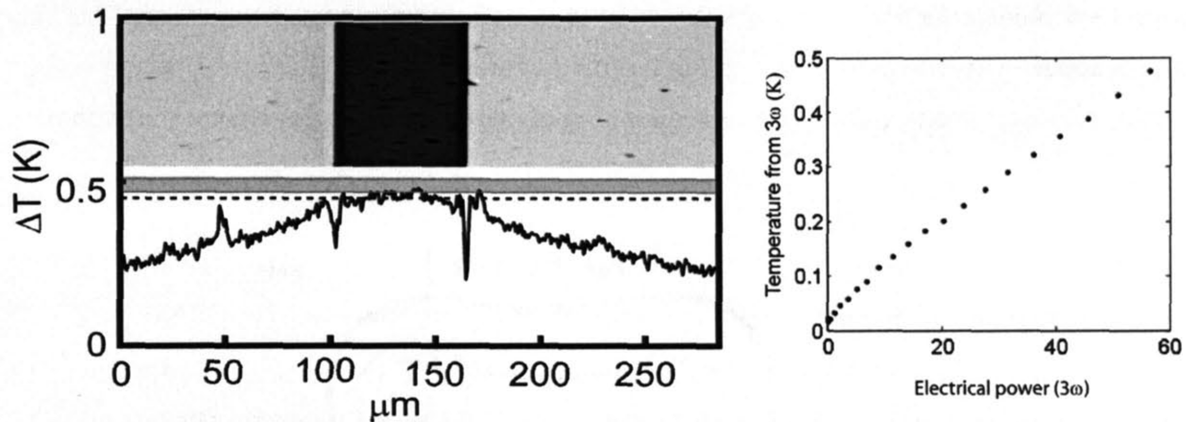


Figure 99. The left-most panel shows a comparison of temperature measured in diffused resistor using a thermorelflectance measurement and using the 3ω technique. The mean thermoreflectance temperature in the resistor is marked with the dotted line and the shaded gray region is the 1-sigma confidence interval for the 3ω temperature. The right panel shows the measured temperature vs bias for the 3ω measurement.

The region of interest for the thermoreflectance measurement is shown in the top of the left-most panel of Figure 99 (the dark area is the heater stripe, with the current running vertically through the stripe). The mean of the thermoreflectance temperature profile (red line) is close (38mK smaller) to the average temperature level found using the 3ω measurement (green line). All of the sources of thermocouple errors considered above would have given rise to a spuriously larger thermoreflectance profile, and so we conclude that they are not occurring here. We also note that the 3ω temperature curve vs input power (rightmost in Figure 99) appears to have a ~15mK offset error in temperature at the origin, most likely due to in-band noise present in the lock-in when the signal is digitally mixed with the reference oscillator. The signal processing origin of this offset error may be identical to that of the thermoreflectance offset error discussed in Section 3.4, and in any case, it may partly explain some of the disagreement between the two techniques. Another source of disagreement between the measurement may be that the temperatures measured with the two methods are not identical, since the thermoreflectance measures the surface temperature and the 3ω measurement can be thought of as measuring a volume averaged temperature within the resistor. We also note that the thermoreflectance measurement and the 3ω technique are performed one after the other rather than simultaneously. This avoids a large electrical nonlinearity that is evidently induced by the green illumination used for the thermoreflectance (for reasons that are not totally clear to us).

It is important to consider any potential sources of error not accounted for in the theory presented above. The question arises: is there a fundamental limit to the accuracy of the technique, provided one is willing to average for long enough to reduce the effects of random noise?

The effects of any slow temperature drifts do not appear directly in the measured thermoreflectance signal, since the lock-in measurement rejects changes in the reflectance occurring at frequencies less that the detection frequency $\omega=2\pi/T$. Drifts may still affect the position of components in the measurement image train, but they are generally easily recognized by the resulting blurring of the image. Such drifts may be mitigated with good environmental control, but are a definite concern for the longer measurements that may be required for high thermal resolution. For certain devices that degrade with thermal cycling, the stability of the device under test may limit the measurement time and the subsequent attainable thermal resolution. We note that drift of the illumination source intensity does not generally adversely affect the measurement because the measured thermoreflectance is the ratio of ΔR to R, both of which would change by the same factor. (In fact, depending on the camera used, slowly varying the illumination somewhat during the measurement might improve the accuracy of the measurement by further reducing any systematic error due to camera nonlinearity.) However, as with other sub-quantization dithered measurements, imperfections in the analog-to-digital (A-to-D) converter may fundamentally limit the accuracy of the CCD thermoreflectance measurements presented here. The errors due to imperfect spacing of the quantization bins of the A-to-D transfer function are traditionally divided into differential non-linearity (DNL) and integral non-linearity (INL) [156]. The DNL describes the difference between the width of a given quantization bin and the ideal quantization bin size, whereas the INL describes the difference of the A-to-D transfer function from an ideal straight-line fit. Both errors are systematic and do not decrease with time-averaging, and so they can potentially distort the temperature signal measured with the camera regardless of how long the measurement is run. Figure 100 shows an example of an imperfect quantization spectrum. The input signal (potentially anywhere on the x-axis) is mapped to a discrete value on the y-axis during the quantization process. The imperfection we refer to lies in the variation of the quantization bin size (the width of the steps in the "staircase") from its ideal size. Shown in the insets of the figure are the size of the DNL error and INL error over 256 quantization bins. Such plots can also generally be found in the specification sheets offered by the manufacturers of A-to-D converters.

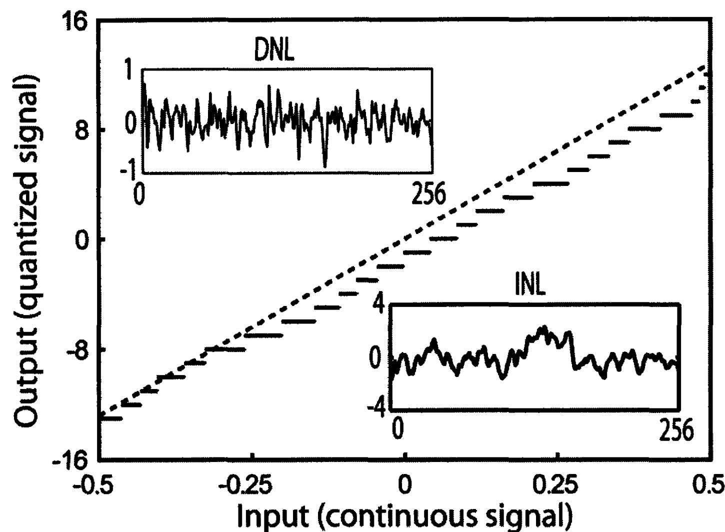


Figure 100. An example of an imperfect A-to-D quantizer. The presence of unequal bin widths simulates the quantizer's nonlinearity error. In this case, a differential nonlinearity (DNL) of 0.35 quantization steps has been assumed. The top inset shows a plot of the DNL error over 256 bins of the A-to-D converter, and the bottom shows a similar plot of the INL error.

For a thermoreflectance measurement where the errors in the quantization levels are locally uncorrelated with one another, the DNL will be the main source of error. According to the manufacturer's specifications, the standard deviation of the DNL error for our camera's A-to-D converter is in the 0.2 to 0.5 range (in units of quantization levels). This might appear to severely limit the effectiveness of the time-averaging relied upon to obtain high accuracy. But once again, the dither present in the signal decreases the importance of these errors. Since many quantization bins will be used to measure the signal, the errors from individual quantization bins are averaged out.

By running a simulated measurement of a small thermal signal with a (randomly generated) imperfect quantizer, we can estimate the effects of imperfect quantization on the accuracy of the measurement. We used a signal that was 5% of the quantization bin with a random noise whose standard deviation was 10 quantization bins in magnitude (on the order of what was measured in our camera). We simulated the lock-in measurement on this signal for 2^{20}=1048576 periods of the signal to reduce any random (statistical) error. The simulation was run 100 times, each with a different imperfect quantizer with a DNL of 0.35 (consistent with the specifications of our camera), in order to get an ensemble average of the systematic error. One instance of the imperfect quantizer is shown in Figure 100. The result of the simulation was that the residual systematic error due to the detector nonlinearity was 4.6% of the signal. For the resistor

measurements shown in Figure 95, this corresponds to a systematic error of 6.7 mK due to A-to-D quantization error.

5.7 Polysilicon-gate field effect transistors

A series of thermal measurements were made on polysilicon gate p-channel FET devices (fabricated in the 6.152J processing class at MIT's Microsystems Technology Lab). The devices were made in a silicon wafer (native doping ~7e15 1/cm^3), and consisted of a 500 nm doped polysilicon gate on top of a 50 nm gate oxide over a boron-diffused p-type channel region. The measurements are presented here because they illustrate the range of phenomena that can be investigated using thermoreflectance, and test the thermal spatial resolution and other measurement limitations. The experimental setup is shown below in Figure 101, along with a photomicrograph of the FET devices.

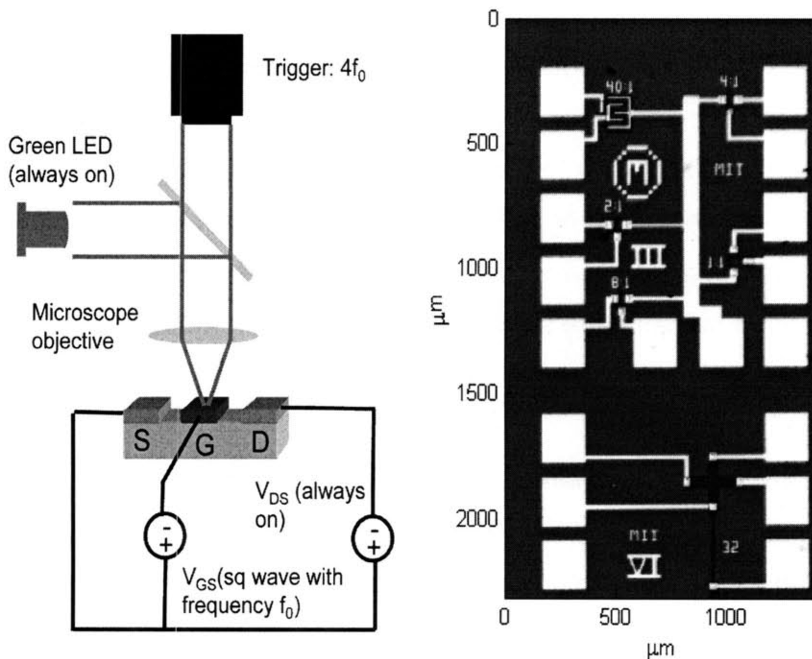


Figure 101. The image at left depicts the experimental setup for the FET thermal measurements. The image at right is a photomicrograph of the wafer with the FETs and other test structures.

The measurements are made by setting a constant voltage bias across the source and drain of the device and switching the gate voltage on and off.

Linear vs. saturation regimes

For the first thermal imaging investigation, we study a large (40 micron gate width, 20 micron gate length) FET, as shown in Figure 102. A 20x Mitutoyo (NA=0.42) was used.

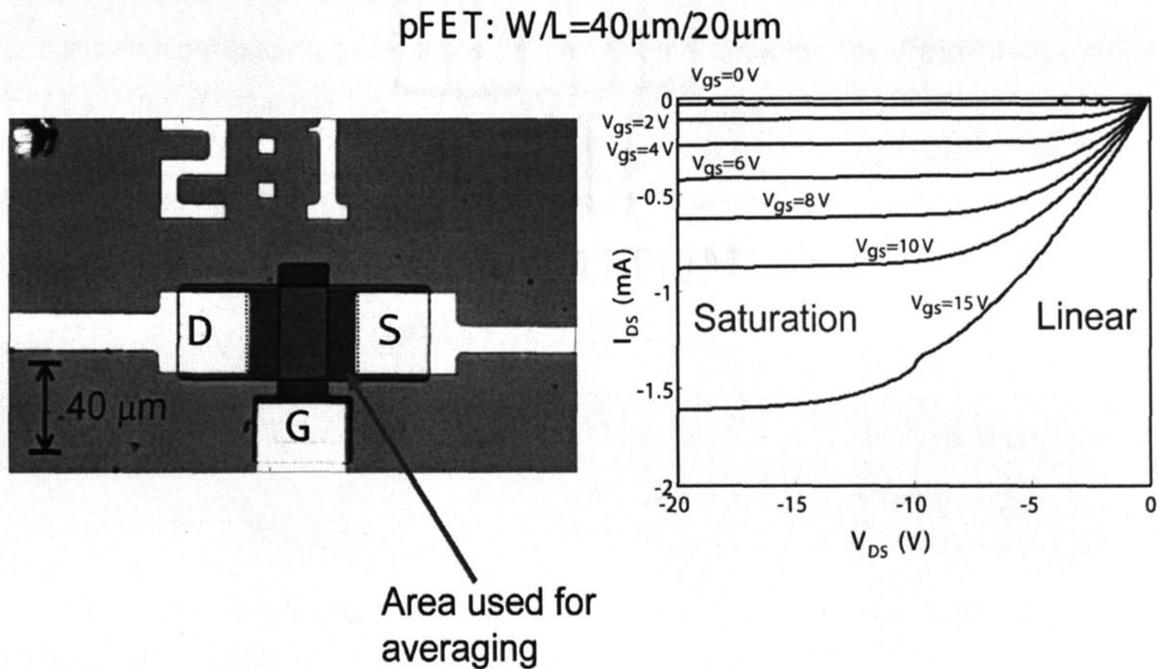


Figure 102. The left-most image is a photo-micrograph showing the p-channel FET to be studies using thermoreflectance, with the the region of interest marked. The I-V characteristics of the FET are shown in the plots on the right. The different curves are for varying gate voltage.

The device is symmetrical between the source and the drain, with the source determined by the side which is grounded. Also shown in Figure 102 is a family of drain current vs. drain voltage curves for varying gate voltage. In the linear regime (low V_{DS}) it is clear that the channel behaves like a resistor whose resistance can be controlled using the gate. The reason is that the charge density in the inversion layer is only slightly perturbed by the applied V_{DS}, making the resistivity under the gate nearly uniform over the gate length. Since the current density in the inversion layer is constant over the gate length, there is a nearly uniform Ohmic ($J^2\rho$) heating profile underneath the gate. In the saturation regime, the V_{DS} is large compared to the gate voltage, and the channel of mobile charge becomes "pinched off" under the drain of the device,

since the local gate-drain voltage there is insufficient to support the inversion of the channel. The drop in mobile charge in the pinch-off region under the drain means that the channel voltage drop is much higher there, so the resulting heating profile has a sharp maximum beneath the drain. In both cases, if we assume the source-drain current density is uniformly distributed under the gate and we neglect any "smearing" of the heat due to local heat spreading, the heating profiles are actually indirectly probing the local electrical potential distribution under the gate.

The results of thermoreflectance measurements on the transistor in both bias regimes are shown below in Figure 103. The measurement was run for 100,000 iterations (2.3 hours per complete image). Also shown are the vertically averaged temperatures in the region defined in Figure 102. The estimated offset error in the thermal images was negligible, and the random error in the gate (standard deviation) was estimated at less than 3%.

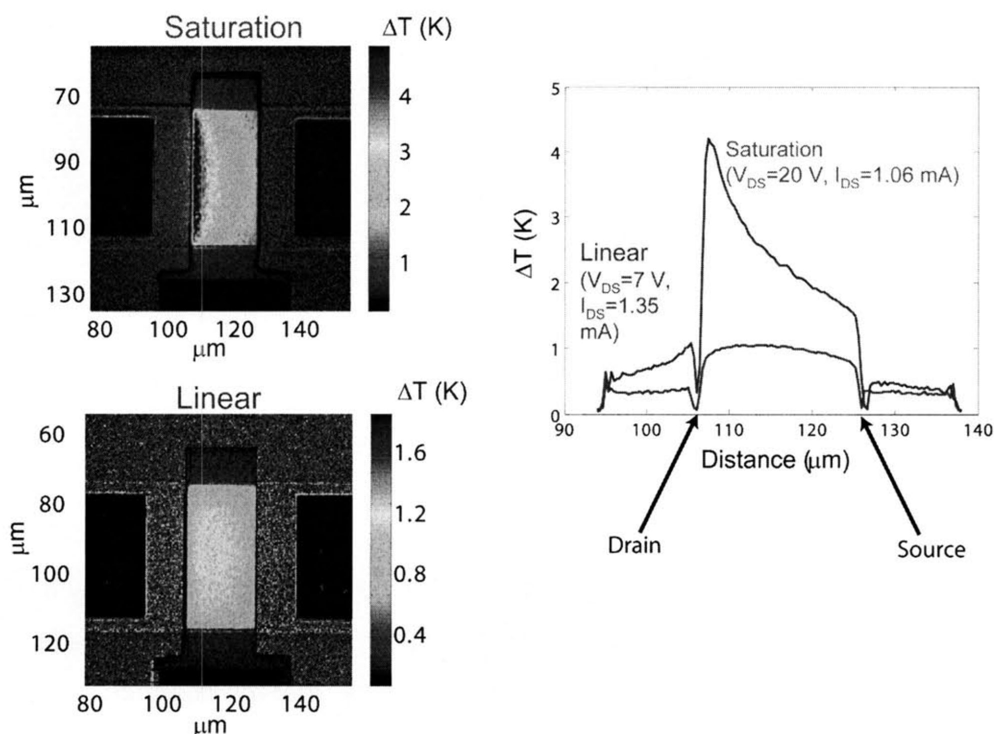


Figure 103. Calibrated thermal images of the FET in the saturation and linear bias regimes. The plots at right show the vertically averaged temperature profiles in the gate (drain is to the left of the gate in the images and profile).

The calibration constants for the doped polysilicon (κ_{TR} =-10.11x10^{-4}) and for the doped silicon channel region exposed to view on either side of the gate (κ_{TR} =2.20x10^{-4}) were used to calibrate the thermal image in those regions. No signal is present in the aluminum contacts (darkest regions in Figure 103) since the camera pixels examining that region were saturated due to the much higher reflectivity of the aluminum relative to the rest of the device.

The saturation and linear regimes can be clearly distinguished according to the sharper heating profile in the pinched off channel. By numerically integrating both of the temperature profiles under the gate we find $|T_{linear}|$=0.96 K and $|T_{saturation}|$=2.45 K, so that the mean temperature in the case of saturation was 2.56 times that of the mean temperature in the linear regime. We can easily compute the dissipated power for each case shown in Figure 103 using the measured V_{DS} and I_{DS}. We find $|Q_{saturation}|$=21.2 mW, $|Q_{linear}|$=9.45 mW, so that the ratio between the two is 2.24. This is close (but not quite equal) to the ratio between the mean temperatures since the thermal resistance seen by the heat in the gate is similar in both cases. In fact, we expect that the thermal resistance seen by the heat in the saturation regime is larger than that seen by the heat in the linear case because the heat is more "bunched up" in the former case, and therefore faces a greater thermal spreading resistance. Thus the increase in mean gate temperature for the saturated case is actually expected to be slightly higher relative to that of the linear case for the same total input power, so that the difference in the ratio of the mean temperatures and the ratio of the total powers is not surprising either. Our implicit assumption that all of the power dissipation in the transistor occurs under the gate could also be a source of discrepancy.

To check the magnitude of the observed thermal signal, we can try to predict the measured temperature in the transistor using the measured power dissipation and an estimate of the thermal resistance seen by the heat. We can estimate the thermal resistance using the spreading resistance model from Chapter 3.3 (the relatively uniform heating in the linear bias regime should approximate the uniform power density assumption implicit in the thermal spreading resistance model). We find that if the area of the gate (40 μm x 20 μm) is used in that model, the spreading resistance seen by the heat leaving the gate is 135 K/W so that an average temperature rise of 1.27 K is expected, compared to the 0.96 K that we measured. This is not an unreasonable disagreement, especially given our implicit (and incorrect) assumption made in this calculation that all of the power dissipation in the transistor takes place under the gate.

One further observation should be made: the thermoreflectance (and hence the temperature) appears to approach zero at the edge of the polysilicon gate. This is not physical; it is an artifact that is caused by the incident probe light scattering off the corner of the polysilicon gate. In general, when the surface of the sample is not flat, artifacts in the thermoreflectance can be caused by light scattering out of the image or by slight movements of the edge due to thermal expansion and contraction of the sample.

<u>High-resolution thermal image of 5 μm gate length FET</u>

To try to get a sense for the attainable thermal resolution for the thermoreflectance measurement we measured the temperature in the smallest device on the test wafer, a 5 μm gate length polysilicon gate FET. We used the highest magnification numerical objective that we have 100x (NA=0.70) for this measurement with 530 nm green light, yielding a 356 nm theoretical optical resolution (using the Sparrow criterion for an Airy disk point-spread function, $r_{min} = \frac{0.47 \cdot \lambda}{NA}$). The thermogram is shown below in Figure 104 for a device biased in saturation:

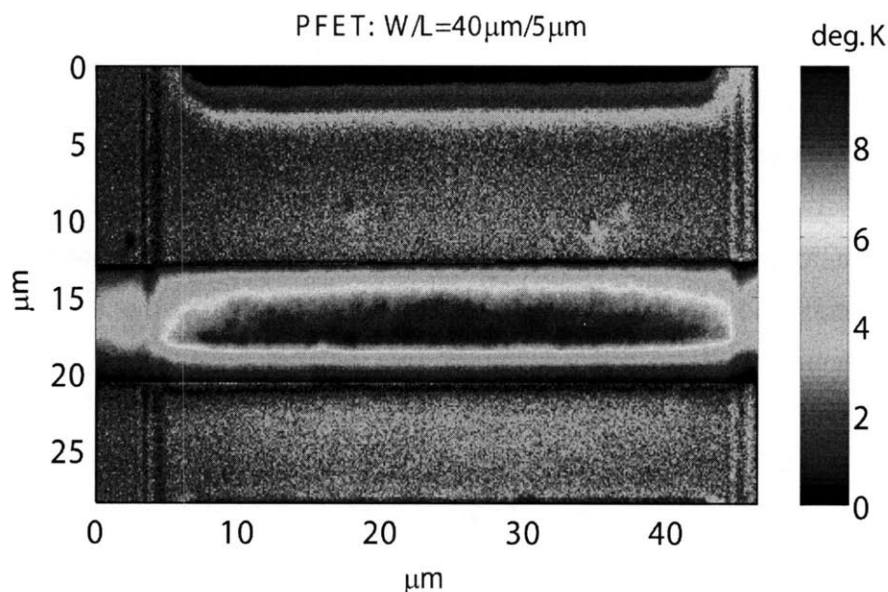


Figure 104. Thermal image of a 5 micron gate-length FET.

It is easy to make out the asymmetrical heating profile, with the largest heating over the drain. Unfortunately it is difficult to quantify the thermal resolution from this image, or from any of the

devices we measured in this thesis. While un upper bound on the optical resolution could presumably be resolved from a deconvolution of the measurement of a sharp edge, to truly measure thermal resolution one needs two areas that have different temperatures but are free from any change in reflectance or surface topography. This is actually a difficult situation to set up. One possible test structure is two long metal lines deposited on an SOI (silicon-on-insulator) wafer, running parallel to one another and spaced apart by 100 nm or less (using e-beam lithography). By using the two wires as resistive heaters that are thermally isolated from one another by the 100 nm gap and the thermally insulating substrate, the two temperatures in the lines could be independently controlled and biased with opposite phase (or different magnitude) Joule heating signals. The resulting phase image from thermoreflectance measurement could be used to establish the true thermal spatial resolution. However, the measurement was not attempted here.

Reliability testing example

The thermal image shown below was taken from a 5-micron gate length FET (Figure 105).

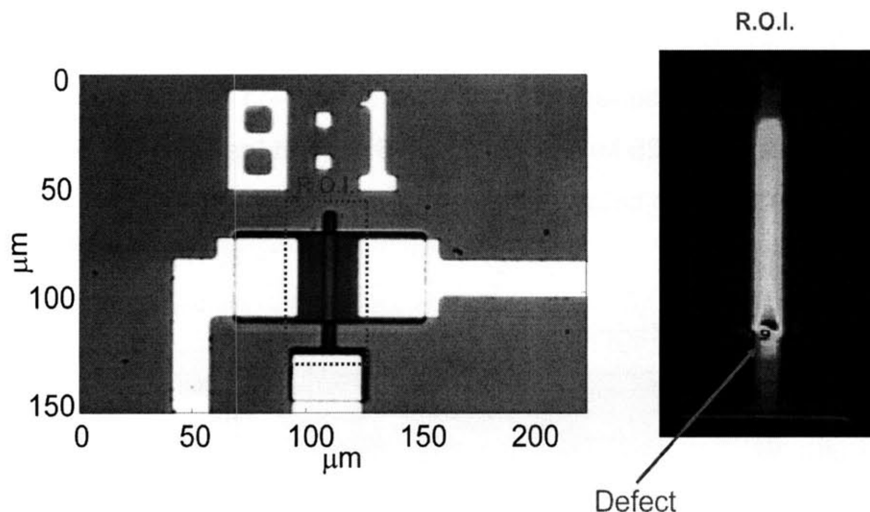


Figure 105. Thermal image of a point defect in the gate of a FET. The defect was not evident in the photomicrograph (shown at left) but was readily apparent in the thermal image and the electrical response of the device.

While there is no obvious defect in the photomicrograph of the device, the electrical characteristics of the device indicated the device was broken. The subsequent thermal image has localized the issue to a point on the gate near the gate contact. The example is intended to

demonstrate the possible application of the thermoreflectance setup to reliability testing of microelectronic devices. There is another possible advantage to this application for thermoreflectance in that the precise magnitude of the temperature is not as critical so that the need for careful calibration (often the most difficult and time consuming part of a measurement) is obviated.

Dynamic heating in a FET

So far the heating in the measured transistors has been due to static power dissipation. In other words, the heating is due to the Joule heating of current passing between the drain and the source with a fixed gate voltage essentially establishing the channel resistance. This is typically the most important source of dissipation in bipolar devices and power transistors. However, dynamic power dissipation is the more important source of power dissipation in many high speed CMOS integrated circuits. It is caused by the charging and discharging of gate capacitances to ground and is therefore highly frequency dependent, becoming important only for high-speed applications. Despite the speed limitation for our homodyne thermoreflectance measurement (camera frame rate <100 Hz), we can still use thermoreflectance to investigate this source of heating. We performed a measurement in which the source and drain of the device were tied to ground, and the gate voltage was rapidly oscillated from 0 to 10V at 25 MHz. The gate voltage was gated with an 8 Hz cycle square wave so over one period of the 8 Hz wave the gate voltage was a rapidly oscillating 25 MHz wave for 62.5 ms and then 0 for 62.5 ms. The results of the measurement are shown below in Figure 106.

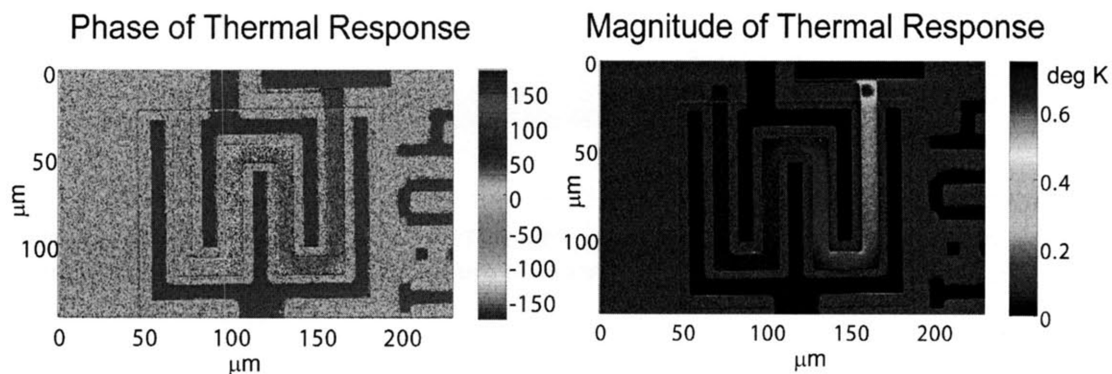


Figure 106. Thermal image of the dynamic heating in a FET gate. The R-C losses in the gate as it is rapdly charged and discharged are responsible for the heating.

Both the phase and magnitude of the measurement are shown. The FET used for this measurement was a fingered structure, and the gate "snaked" through the structure between the source and drain metal. Both the phase and magnitude have a heating profile consistent with the dynamic heating of the gate polysilicon. The polysilicon gate is not perfectly conductive so as the gate is successively charged and discharged, there is heating along the length of the gate. The heating grows in magnitude and the phase signal becomes more pronounced near the gate contact (in the upper right hand quadrant of the images) because the closer to the contact a greater fraction of the gate charge passes through the polysilicon due to the distributed charge storage. There are two other features worthy of comment. The first is the square of intense heating evident in the magnitude image in the polysilicon gate next to the contact. This was initially mysterious, but a check of the mask used for the doping implant of the polysilicon revealed that (for reasons unknown) that part of the gate was left undoped. Its resistance is therefore higher, and the bulk of the heating occurs there. The other feature is a spurious effect that can be most clearly seen in the phase image. The silicon substrate should not have any significant thermal signal, and therefore we would expect the phase image in the substrate to be noise. Instead, it has a fairly well-defined phase. This was actually tracked down to be due to spurious electromagnetic coupling between the unshielded dc-probes and the bias circuit for the (constant intensity) LED source used for illumination in the measurement. It can be reduced by using coaxial probe tips and good grounding practices.

Before moving on, we will address a potentially important issue that has been ignored so far during the FET measurements. All of the thermoreflectance measurements in this thesis depend on thermal modulation. It is important to remember that for the measurements to be interpreted as the temperature profile of a device under constant bias (or as the mean temperature profile of a device under a constant square wave bias, in the case of the dynamic heating) the devices must "thermally charge" and relax in a time scale that is short compared to the shutter time of the CCD camera. This way, when the reflectance of the sample (and hence its temperature) is sampled by the camera, it reflects the steady state operation temperature of the device. This time scale for the FET devices is easy to estimate using the dynamic heat spreading theory discussed in Chapter 3. We calculate that for a 40 μm x 20 μm device on silicon, the thermal resistance from the device to thermal ground changes by less than 1% between dc conditions and a time scale of 10 ms. For all of the measurements in silicon wafers made for this thesis, the error from this effect should be negligible. Also, the measurement can

always be slowed down to minimize this effect. A measurement (not investigated in this thesis) for which the dynamic thermal resistance would actually be useful could be made by applying a rapid (or sinusoidal) thermal pulse to the back side of a (polished) sample. By measuring the response time (or phase lag) in the thermal response on the polished front face of the sample using a thermoreflectance phase locked to the source, the diffusivity (or conductivity, if the heat capacity is known) could be extracted.

5.8 Cross-plane thermoreflectance imaging of Bi_2Te_3

The thermoreflectance technique was applied to single n-type polycrystalline Bi_2Te_3 thermoelectric elements, 508 μm thick and 1 mm^2 in area. Much of this work was presented at the Fall 2005 MRS meeting[157]. Previous work by others has established thermoreflectance imaging as a tool for characterizing thermoelectric micro-refrigerators by measuring the temperature change in a metal contact on the device [132,137], or by measuring light's thermal-expansion-induced deflection off of a heated TE material[140]. We extend that work here by thermally imaging the most important TE material in the cross-plane direction to directly observe the heat transport through the element. By biasing the element with a sinusoidal current and locking in successively on the Peltier and Joule components of the heating, the contributions to the heating from each can be separately observed in a thermoreflectance image. Using a calibration technique, the image can then be converted to a temperature image of the device. A finite-difference solution can be applied to compare the operation of the device with theory.

The thermoreflectance measurement setups used here are shown schematically in Figure 107. In both the measurement (left) and calibration (right) a steady source of white light is used to illuminate the sample. The temperature of the sample in the measurement is modulated using a sinusoidal current source of frequency ω. The CCD camera then takes pictures of the sample at a frequency phase-locked to the excitation. To detect the amplitude of the Peltier heating and cooling occurring at a frequency of ω, the camera takes images of the device at a trigger frequency of 4ω. To detect the amplitude of Joule (Ohmic) heating, the camera takes images of the device at a trigger frequency of 8ω, since the temperature change from this heating (I^2R) is at 2ω. The main quantity of interest is the magnitude $\left|\frac{\Delta R}{R}\right|$.

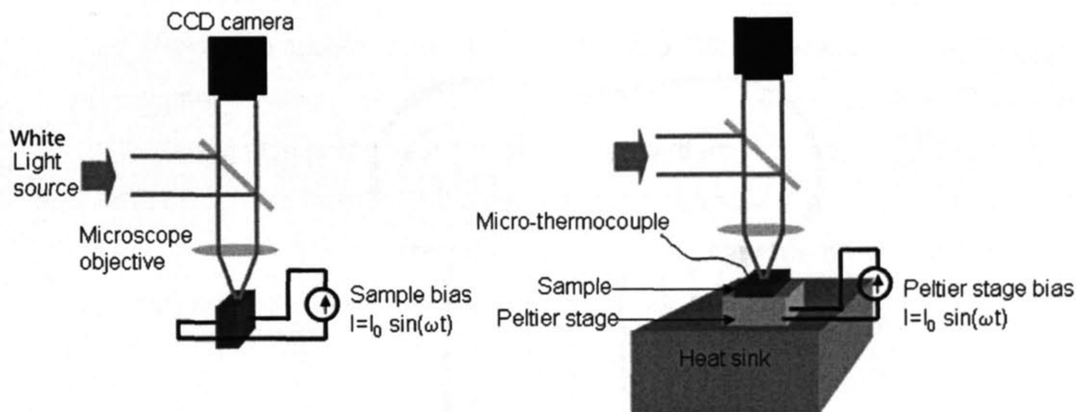


Figure 107. Schematic diagram of the thermoreflectance imaging apparatus used for measurement of thermoelectric elements (left) and calibration of the material thermoreflectance coefficient (right).

The calibration setup is shown on the right of , and is identical to the technique outlined in the previous section. A flat, uniform sample of the material of interest is placed on a high-speed Peltier cooler. The top contact has been removed with polishing, exposing the smoothed polycrystalline Bi_2Te_3. The temperature of the Peltier stage (and sample) is modulated using a sinusoidal current. A thermoreflectance image of the sample is obtained in the usual manner, and simultaneously the temperature on the top of the sample is measured with a 25 μm diameter microthermocouple attached with heat-sink compound. A low numerical aperture lens (0.28) is used for the calibration to avoid any parasitic effects due to sample motion from thermal expansion of the Peltier cooler[158]. Although the sample has been rotated between the measurement and calibrations, this should not be important in a polycrystalline material with random grain orientations.

In order to obtain a good thermogram a flat sample surface is required since scratches and other defects can move slightly with the thermal expansion and contraction of the material, adding spurious noise to the thermoreflectance image. A 508 μm thick, 1 mm^2 area sample of polycrystalline Bi_2Te_3 was edge-polished using a series of aluminum oxide abrasive slurries from 3 μm to 0.05 μm in mean particle diameter. A separate element from the same lot was face-polished for the thermoreflectance coefficient measurement. A short nap polishing cloth is used to avoid edge rounding, since the Bi_2Te_3 is rather soft. Despite these steps, surface quality was somewhat poor compared to typical polished semiconductor samples because debris from the polycrystalline Bi_2Te_3 can break off and scratch the face of the sample. The

calibration measurement was performed as described in the previous section, and a pixel average over the (uniform) images was used to calculate the thermoreflectance. A thermoreflectance coefficient of κ_{TR} =(3.78 +/- 0.05) x 10^{-4} 1/K was found using a temperature modulation of amplitude 1.07 K. To confirm that sample motion was not affecting the measurement, the measurement was checked again at double the magnification, with the same results. This is the first measurement of the thermoreflectance coefficient for polycrystalline Bi_2Te_3 that we know of, although single crystal samples have been measured in the past, as shown in Table 6 [159].

Material	Illumination	Measured κ_{TR} (mag.)
Gold	Blue	3.29 x 10^{-4}
Silicon	Green	1.36 x 10^{-4}
Polycrystalline Silicon	Green	10.1 x 10^{-4}
Crystalline Bi_2Te_3	Blue	1.5 x 10^{-3}
Polycrystalline Bi_2Te_3	White (broad)	3.78 x 10^{-4}

Table 6. Measured thermoreflectance coefficients for various important materials.

A broad white light source (above the band-gap[160]) was used since the thermoreflectance spectrum of the specific material was unknown and polarization-specific measurements in the literature on cleaved single crystal Bridgeman-grown Bi_2Te_3 suggest a large negative thermoreflectance coefficient over the visible light regime. Shown in the Table 6 are some measured thermoreflectance coefficients determined using this calibration technique[161]. Only the magnitude was determined from thermoreflectance because there was no other known phase reference.

Once a satisfactory finish was obtained for the edge polished sample, it was ultrasonically cleaned in water and mounted between two spring-loaded copper contacts. A 2A amplitude sinusoidal bias current was used at two measurement frequencies (0.5 Hz and 7 Hz). Two different styles of electrical and thermal contacts were tested, one with thin (~127 μm) copper sheet contacts and one with thick (~2 cm) copper bar contacts. For all measurements, 10^5 iterations were performed.

The resulting images were averaged in the transverse direction (relative to the current's direction) in order to improve the temperature resolution and obtain a thermal profile of the element along the direction of current. For all samples and bias conditions, the temperature profiles of the Peltier effect displayed a clear maximum near the interfaces with the metal contacts, while the effects of Joule heating were largest inside the sample. Also, the Peltier phase profile displays a characteristic phase shift of 180 degrees across the sample, since one side is heated as the other is cooled. This effect is plainly visible in the phase image, an example of which is shown in Figure 108. In contrast, the Ohmic phase is constant in sign, since the Ohmic temperature changes are always caused by heating. (Both phase profiles average to zero when the signal becomes too small or when undergoing a switch across the sample). The noise in the temperature profiles is not set by the thermoreflectance, but rather by the microcrystalline structure of the sample. The "speckle" visible in the figure is consistent with 10-20 μm particles, flattened in the transverse direction, just as we would expect the microcrystalline material to look if it were produced used a pressed-powder metallurgical manufacturing process (as discussed in Chapter 4).

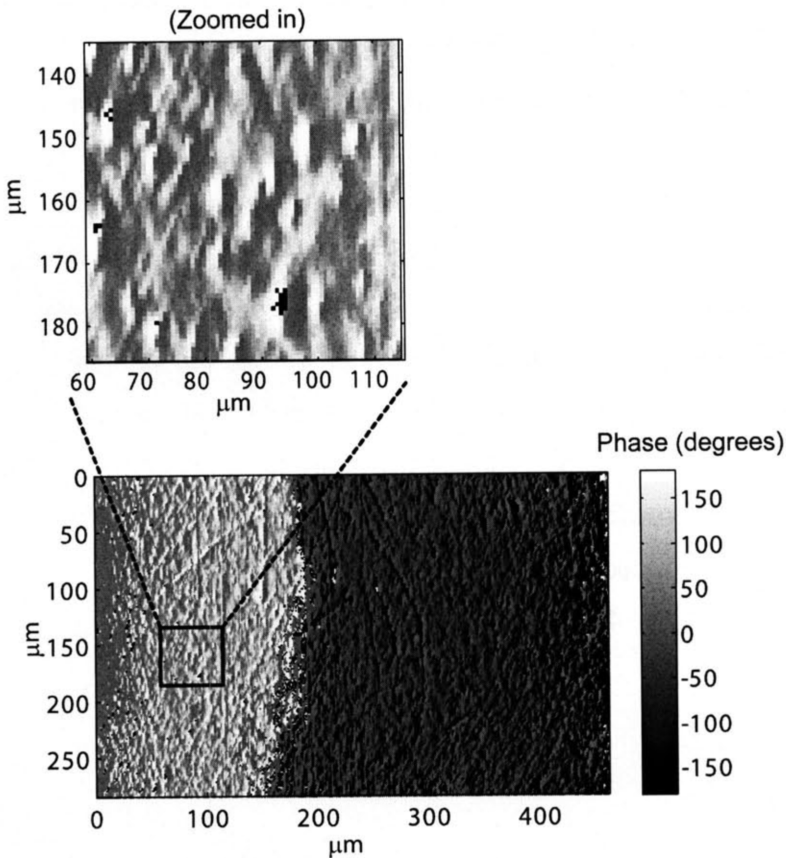


Figure 108. Phase image from Peltier measurement of Bi_2Te_3 sample. The speckle in the phase image is not visible in the photomicrograph of the sample, and the size of the speckle matches the expected polycrystalline grain size (~10-20 μm).

Shown below are the Peltier and Ohmic temperature (and phase) profiles of the Bi_2Te_3 measured with the slow (0.5 Hz) current excitation and the thin copper contacts (Figure 109). We note that many of the expected features are visible. The Peltier phase does undergo the expected reversal in the bulk of the sample while the Ohmic phase profile is of a single polarity. However, the phase image for the Peltier effect at the edge of the sample near the contacts appears to lag behind the phase on the inside of the sample.

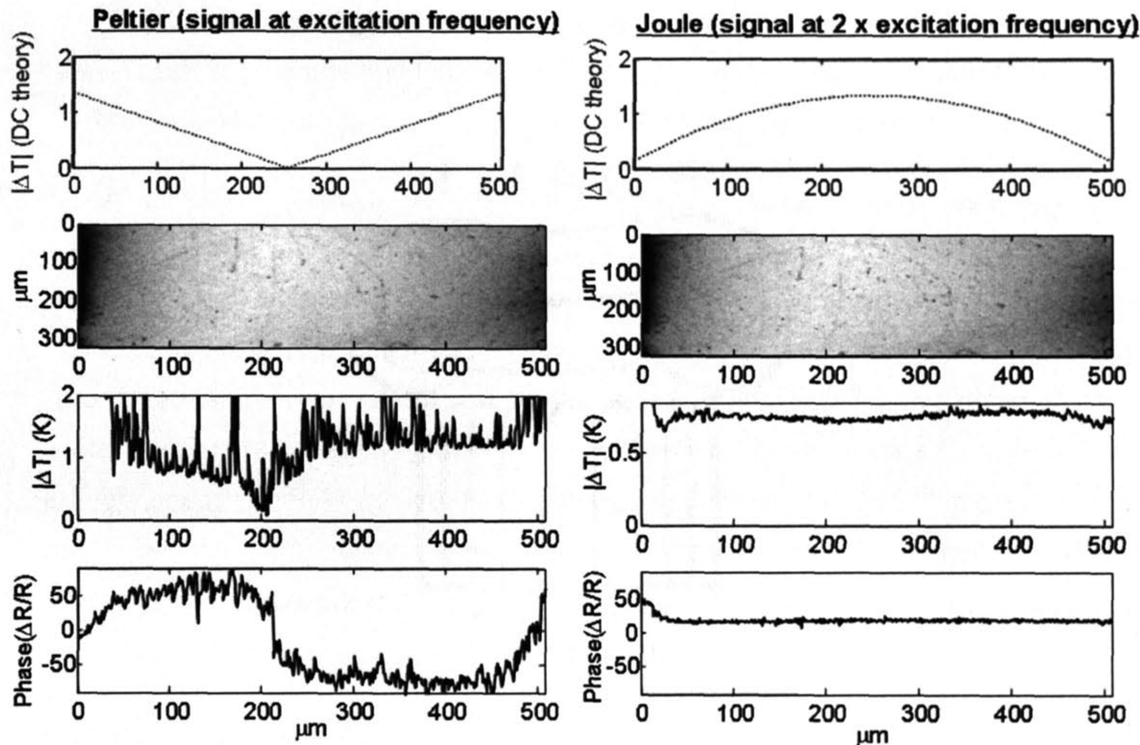


Figure 109. Cross-plane thermoreflectance images of Bi_2Te_3. The left most images depict the Peltier measurement (solid) and steady-state theory (dotted), and the right depicts the Ohmic. These measurements were made using thin copper thermal and electrical contacts and a 2A sinusoidal current at 0.5 Hz.

Also shown in Figure 109 (dotted top curves) are some predicted thermal profiles. The predicted profiles were obtained using the steady state finite-difference Onsager solver discussed in Chapter 2.6. Using material parameters obtained from a larger TE element made of the same material (Appendix 3 [162]) by Marlow Industries, the measured temperature profiles can be fit by assuming steady-state conditions. A Seebeck coefficient of -240 μV/K, a thermal conductivity of 0.0147 W/cm/K, and an electrical conductivity of 737 1/Ω/cm were assumed, corresponding to a thermoelectric figure of merit ZT=0.89 (see Appendix 3). Also required for the simulation is the effective heat sink parameter h_{eff} describing the thermal interface with the copper contacts. This parameter has units of $W/cm^2/K$ and is assumed for simplicity to be symmetrical and fit to match the magnitude of the Peltier measurement (h_{eff} =20 $W/cm^2/K$), although it seems evident from the position of the abrupt phase shift in the material that the thermal contacts to the sample were slightly unbalanced.

One reason for the difference between the measured temperatures and the steady state simulations is that the simulations do not include dynamic effects taking into account the external thermal capacitance of the copper contacts or the internal heat capacity of the Bi_2Te_3. The latter may be important due to the low thermal diffusivity D of crystalline Bi_2Te_3, around 0.011 cm^2/s [77]. Since the thermal penetration depth in the device is given by $\lambda = \left|\sqrt{\frac{D}{i2\omega}}\right|$ (from [57]), we would expect a penetration depth on the order of 125 or 175 μm for the Joule or Peltier heating, respectively. The thermal penetration depth may be even smaller due to the fact that the polycrystalline sample has a lower thermal conductivity than the crystalline data used here for D. It may be due to the short thermal penetration depth that the phase of the Peltier profile varies rather slowly from 90 degrees (heating) to -90 degrees (cooling). Another reason for the discrepancies is that the current is not uniform in the averaged dimension (as evidenced by the curved phase boundary in Figure 108. This causes a "rounding off" of the sharp 90 degree phase transition that is ideally expected across the material. The speckle due to the polycrystalline sample also causes some noise in the profiles.

Some of these inferences can be tested by changing the frequency and bias conditions of the measurement. A measurement taken using a 7 Hz current signal (14 times faster) but otherwise identical to that of Figure 109 is shown in Figure 110. The results show significant decreases in the magnitudes of the observed thermal signals, confirming that time domain effects are important in interpreting the profiles.

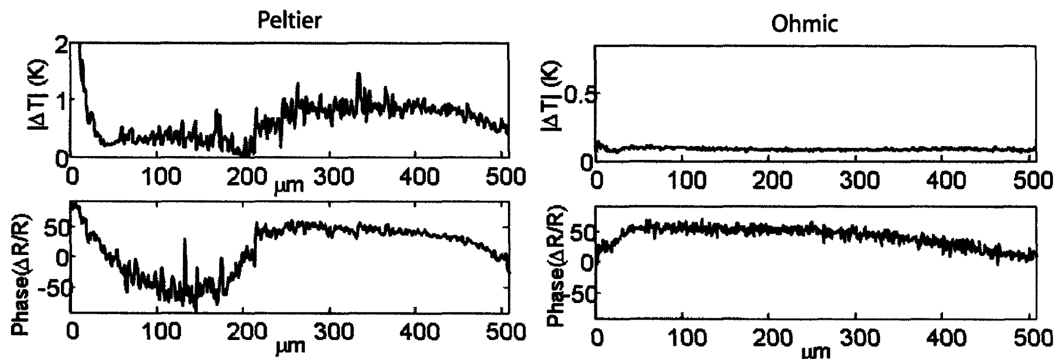


Figure 110. Peltier (left) and Ohmic (right) images demonstrating the importance of the frequency of excitation. These measurements were made using thin copper thermal and electrical contacts and a 2A sinusoidal current at 7 Hz.

This study demonstrates that the thermoreflectance imaging technique can be used to othe temperature distribution in a thick-film Peltier element in the cross-plane direction by separately locking in on the Joule and Peltier components of the temperature distribution. However, while the measured temperatures have the general behavior expected from a simple one dimensional model, quantitative agreement between the present theoretical model and the data is not possible using the steady state model because time-dependent thermal effects have been neglected. Recently, there has been work presented by Downey et al. on a transmission-line style modeling of thermoelectric couples[163] that appears to capture the necessary time dependence. We have attempted our own solution by writing a one-dimensional time finite difference time domain solver using a Crank-Nicolson discretization that is accurate to second order in both time and space and is also unconditionally stable. The equations for heat transport we solved were the time-dependent Onsager relations:

$$C\frac{\partial T}{\partial t} = \nabla\cdot(\kappa\nabla T) + \frac{J^2}{\sigma} - JT\nabla\alpha$$
$$E = \frac{J}{\sigma} + \alpha\nabla T \tag{5.7.1}$$

The thermal boundary condition assumed on either side of the thermoelectric was a complex thermal impedance that could be used to model an arbitrary thermal boundary condition at the excitation frequency. The impedance for these simulations was set to match the contact resistance measured in Chapter 4 for the Bi_2Te_3 under similar contact conditions (0.1 K cm^2/W) in parallel with a thermal capacitance equal to the approximate thermal mass of the copper tabs used to contact the devices (0.10 J/K/cm^2). The properties of the MAM Bi_2Te_3 were once again assumed to be exactly those measured in the Marlow characterization of Appendix 3, with a diffusivity consistent with the book estimate for Bi_2Te_3 (0.011 cm^2/s). The solver was capable of predicting the voltage and thermal response to an arbitrary applied current but for the cross-plane thermoreflectance study a sinusoidal current density matching that used in the measurements was used (J=200 A/cm^2). The only "fitting" parameter we used was to increase the effective heat transfer coefficient of the left hand side of the device by 50% in order to demonstrate how a mismatch in the thermal contact resistance could cause the "V" shape of the Peltier thermal response to shift its position (and more closely resemble the measured profile). The results of the simulation are shown below for both the "low" frequency (0.5 Hz) and "high" frequency (7 Hz) current excitations.

Some of the features in the simulations match the measured results better than the steady state thermal response shown in Figure 109. The magnitude of the thermal response at higher frequencies is observed to decrease, as we would expect. The phase response of the Ohmic heating is beginning to display some structure at the higher frequencies due to the low thermal diffusivity of the Bi_2Te_3. The phase change of the Peltier effect in the middle of the sample has been smoothed somewhat from the idealized picture, similar to the measured response. But in general the fit is not good. In particular, the peculiar features at the edges of the sample (like the sharp changes in the observed Peltier response) do not appear to be well-explained.

The Fourier transforms of the thermal response that we plot here are unfortunately very sensitive to the particular values chosen for the estimated thermal boundary conditions and the estimated thermal diffusivity of the MAM Bi_2Te_3, making the prospects for using the cross-plane thermoreflectance measurement for quantitative parameter extraction somewhat dimmer. Part of this is due to the fact that the relevant thermal time scales in the problem are rather close to the measured frequencies: The thermal RC of the contacts is around 1 s, and the characteristic time defined by the internal heat capacity of the Bi_2Te_3 is around 0.2 s, so that errors in these time scales will cause relatively large alterations in the simulated profiles. It is possible that by going to low frequencies some of the uncertainties may be mitigated, but due to the inherently poor quality of the image due to the polycrystalline samples, this was not pursued further. Higher frequencies might also be examined, but since the maximum frame rate of the camera is around 100 Hz, the maximum Ohmic heating frequency that can be measured with a homodyne setup is around 100/4/2=12.5 Hz.

5.9 Conclusions

In this chapter we have presented measurements made using the technique of thermoreflectance lock-in CCD imaging, a non-contact means of thermography capable of sub-micron resolution. We have presented an analytic mathematical model of the measurement that allows the attainable resolution of the measurement to be quantified by propagating fundamental sources of noise through the signal flow of the measurement. This model was captured in a single analytic function that returns error bars for the measurement given the known measurement conditions and two parameters describing the noise in the camera. We have described and demonstrated a method for characterizing the noise in a camera, and tested the resulting experimental error bars against our theoretical error models. A method for

calibrating the thermoreflectance coefficients for an arbitrary sample was presented and demonstrated. We have demonstrated a thermal resolution of 18 mK and checked the temperature against an independent measurement using a 3-ω setup built for that purpose.

We have investigated the practice of micro/nano-scale thermoreflectance imaging with a series of measurements on polysilicon gate FET devices and a MAM Bi_2Te_3 thermoelectric element. The former were used to investigate the potential of the technique for the imaging of active semiconductor devices. With the exception of a few artifacts that we have discussed (sample tilting/motion due to thermal expansion, rough surfaces/edges, electrical interference with the probe light) these temperature images measurements conformed with our theoretical expectations and physical intuition. That latter measurement demonstrated the first thermoreflectance imaging a thermoelectric element in the cross-plane direction. While quantitative measurements of the temperature were possible, modeling of the device using a finite-difference time domain solver was complicated by the sensitivity of the measurement to thermal boundary conditions whose magnitude could only be estimated, and to the multicrystalline character of the sample. We judged that the match between experiment and theory in this case was not yet good enough for quantitative thermoelectric property characterization.

Chapter 6: Conclusions and Future Work

This thesis has presented our work and thinking on the problem of high density thermoelectric power generation and the related issues of thermoelectric metrology and non-contact sub-micron thermal imaging. We review the main conclusions of this work here, and also suggest some areas for future exploration.

Research efforts directed towards improving thermoelectric power generation and cooling have traditionally been directed towards improving the material figure of merit, $ZT=\alpha^2\sigma/\kappa$. The material ZT sets a theoretical upper limit to the efficiency of thermoelectric power generation and coefficient of performance for cooling. Also, because it is the simple combination of three fundamental material parameters (and temperature), the figure of merit has proven to be a very useful coordinating device for addressing the materials science challenge associated with improving the performance of thermoelectrics. In this thesis we have approached the problem of thermoelectric power generation from a different direction, from an engineering standpoint rather than from a materials science standpoint. This has given us a slightly different perspective on thermoelectric power generation general, and high power density (thin-film) generation in particular. We summarize the essential reason for this shift in perspective in Figure 111.

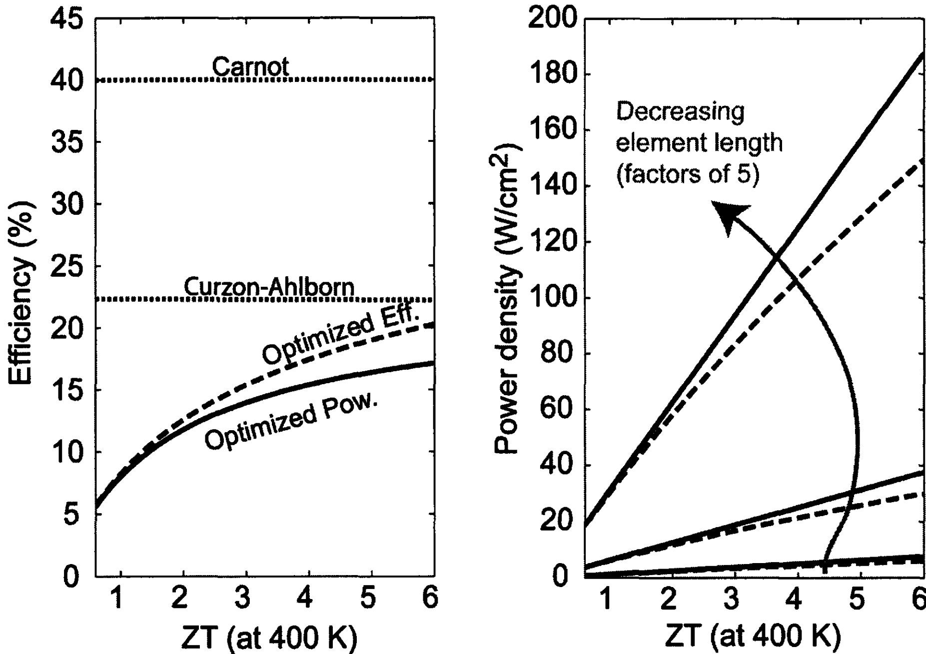


Figure 111. The left plot plots the efficiency optimized for power (solid) and the efficiency optimized for efficiency (dashed) as a function of ZT. The right plot does the same for generated power density (optimized for power = solid, optimized for efficiency = dashed) for three element lengths (2mm, 400 μm, 80 μm)

The plot depicts the expected performance of power generating thermoelectric elements as the material ZT is increased. The left-hand plot shows the maximum possible efficiency for two different optimization criteria, maximum power and maximum efficiency. Also plotted are two ultimate limits to the efficiency, the Curzon-Ahlborn limit (for the optimized power) and the Carnot efficiency (for the optimized efficiency). It is striking that improvements in efficiency come frustratingly slowly with increasing material ZT, and (for the example shown here) even for ZT=5 we are a long way from the Carnot efficiency. Contrasted with this picture is the plot of generated power density in the right hand part of the figure. Plotted again are the two optimization criteria, and also shown is the effect of decreasing the length of the thermoelectric element by factors of 5, from an element length of 2 mm (bottom pair of curves), to an element length of 400 μm (middle pair of curves, to an element length of 80 μm (top pair of curves). The optimal generator power density increases rapidly with decreasing length, especially if ZT is also improved.

We were motivated by this picture and by the current range of thermoelectric power generation efficiencies (0.1 W/cm^2 – 5 W/cm^2) to try to seek performance improvement in power density as well as the ZT (efficiency). We also knew that technology for handling power densities in the 100's of W/cm^2 have been developed for cooling microprocessors and laser diode arrays. Neglected in Figure 111 were the limitations of this approach, namely parasitic electrical resistance and the finite rate heat transfer in and out of the generator mandated by the heat sink and any parasitic series thermal resistances. Chapter 2 and Chapter 3 of this thesis were devoted to investigating these limits, and the ultimate feasibility of high power density thermoelectric generation.

There were several important conclusions we drew from this investigation. Limitations on the maximum heat transfer coefficient of presently available heat sinks and on the thickness of nanostructured TE elements force present thin-film generators into a heat sink limited regime, characterized by a large temperature drop across the heat sink and across parasitic thermal resistances compared with the temperature drop across the device. In this regime, we found that some of the conventional wisdom of thermoelectric generator optimization must be revised. We showed that the optimum operating current of the generator becomes smaller, so that the normal impedance matching condition guaranteeing maximum power no longer holds, and we derived the condition for the new optimum. Also, the relative impact of the three principal TE material parameters (α, σ, κ) on generator performance is altered in the heat-sink limited regime. Perhaps most surprising was that maximizing the thermoelectric figure of merit Z no longer guarantees optimal efficiency. In the heat sink limited regime changes in the thermal conductivity are the most important for the generated power density; this differs dramatically from the ideal case where the generated power density is insensitive to changes in thermal conductivity.

Having motivated power generation as a design goal and having identified the key obstacles to improving the generator performance, we proposed an idea for a solution using optimized heat spreading and a microchannel cooler to remove heat from the cold side of the generator. To model this we developed tools to describe heat and electrical transport through each of the components, and combined them to calculate the net generator performance. We concluded that even using an unexceptional thermoelectric material, substantial improvements to the generated power were possible over both a conventional thermoelectric generator (120x improvement) and an optimized generator cooled using an aggressive passive heat sink (6x improvement). These improvements came without correspondingly large decreases in

efficiency ($\Delta\eta$=-0.45% compared with idealized passive model, and $\Delta\eta$=-0.75% compared to the commercial generator.)

We would like to suggest a few fruitful areas for future work in the areas of this field discussed in Chapter 2 and Chapter 3. Our results have suggested that a properly optimized high-power density generator with outstanding performance could be physically realized using conventional thermoelectric material (ZT~1). With our theory of Chapter 2 and our concrete model in Chapter 3, we have tried to show that attention to parasitic elements and heat transfer is crucial for good design in this space. We think that the some of the heat management technologies being pioneered for the semiconductor industry (we chose microchannels for our study in Chapter 3) will also be good solutions for compact, high power thermoelectric generators. While we have made the first steps towards that goal with our modeling, the construction and testing of such a system may reveal other important issues that we have not introduced into our model. An intriguing idea for the use of a microchannel cooled TE generator is waste-heat reclamation on the exhaust pipe, radiator, or engine of an automobile. The large and growing market for automobiles and the (hopefully) ever-increasing importance of fuel efficiency are good motivation for work in this area. Such an application demands high power density from a compact system and also already has much of the overhead (radiator, water cooling loops, pump, fans) that might be needed to run the microchannel cooling loop. A serious unresolved issue in applying the model of Chapter 3 to this design is the engineering of the hot side heat exchanger to remove the heat from exhaust gas, or radiator water. Many of the design choices inherent in such a discussion are highly specific to the operation of a car (e.g. effect of gas back pressure due to the flow resistance of the hot side heat exchanger of the generator on the engine or on the catalytic converter), and so we did not investigate them here. Another interesting use for the microchannel cold plate TE system might be for the characterization of power generation in TE elements. As we have stressed, the same issues that are associated with finite heat transfer affect both the practical use of thermoelectric thin films and their characterization. A test bench made from two such microchannel heat exchangers (one for the hot side, one for the cold side) could more closely approximate the ideal temperature boundary conditions desired for thin-film TE element testing than any test system we are aware of.

Having discussed the important system level issues in Chapters 2 and 3, in Chapter 4 we designed and tested high-temperature test stations for thermoelectric real-conditions device testing. The need for this was motivated by a discussion of current thermoelectric thin-film

metrology. The heat transport and parasitic limitations that motivated the microchannel system design in Chapter 3 are also a serious issue for characterization of the films themselves. We believe that as a result, there are unresolved issues with many of the thin-film thermoelectric measurements in the literature, including the result of ZT>2 using the transient Harman technique[22] and the result of ZT>1.3 using the maximum cooling temperature (and related techniques)[21].

Using these tools, we measured thermal conductivities, Seebeck coefficients, power, and efficiency in elements ranging from bulk commercial elements to epitaxial thin-films. Our measurements of the Seebeck coefficient and the thermal conductivity of the MAM Bi_2Te_3, a state-of-the-art commercial material, agreed with those provided by the maker of the material (Marlow Industries) up to a measurement error due mainly to the parasitic thermal interface resistance in our measurement. Subsequent measurements on an array of different length elements conformed with our theoretical expectations, with the power increasing as the elements were scaled to shorter lengths until parasitic electrical lead resistance began to dominate. We showed that using the measurements on single elements it is possible to infer the performance of a packaged generator with reasonable accuracy, with the exception of an error in the electrical resistance (and quantities derived from it) due to the aforementioned parasitic electrical lead resistance.

Our measurement of the Seebeck coefficient of the Lincoln Labs G-207 PbSeTe/PbTe (also known as n-QDSL)[21] was α=-216 μV/K, at a mean element temperature of 347 K. This was within the range of measured Seebeck coefficients reported by Harman et al. (-208 μV/K - 219 μV/K at room temperature[21], and 170 μV/K – 240 μV/K at room temperature for a larger sample set elsewhere[164]). Our measured thermal conductivity for this material was between 8.8 mW/cm/K and 9.5 mW/cm/K for a mean element temperature of 347 K, with the lower number being what we directly measured, and the larger number resulting from a subtraction of the interface contact resistance, which we estimated at 8% of the total thermal resistance of the element. (Because the interface conductivity was the dominant error in the measurement, we conclude that the true material thermal conductivity is strictly greater than 8.8 $mW/cm^2/K$.) This disagreed with the thermal conductivities reported by Harman at al.[21] We have pointed out that the Harman thermal conductivity result appears to be based in part on an incorrect usage of a formula for the thermoelectric figure of merit, calling into question the accuracy of their thermal conductivity and all subsequent reports of ZT>1 based on that thermal conductivity.

We also measured the open circuit voltage from a thin ErAs nanodot film (5 μm) on a family of differing substrate thicknesses, so that the effects of the substrate could be removed. The results were compared with the expected open circuit voltage using other measured values for the substrate and thin-film thermal conductivities and Seebeck coefficients. Our measured open circuit voltage was consistent with the expected open circuit voltage for the longer length substrates (<10% error) but began to show substantial disagreement for the shorter substrates (~25% error). We conclude that the uncertainties resulting from the size and irreproducibility of the thermal contact resistance of the elements severely limited the accuracy of the technique. In general, the measurements became progressively more difficult for the thinner elements as we pushed the limits of applicability of the macroscopic heat and electrical measurements. Electrical and thermal parasitics eventually limited the accuracy and reproducibility of the measurement, motivating us to investigate an alternate approach to thermal metrology that could avoid some of these problems.

There is also a need for future work on a more quantitative theory that can be used model the ZT of the ErAs nanodot metal superlattice system, in particular any thermionic effects that may be present. The concept of the Peltier coefficient is clear to us when the conditions assumed during its definition (through the Onsager relations of first-order non-equilibrium thermodynamics[40]) or through its more modern derivation (in the low-field limit of the Boltzmann transport equation[2]) are met. There is also a sense in which a "local" Peltier coefficient can be defined. One begins with a definition of the average energy of the current distribution[165]:

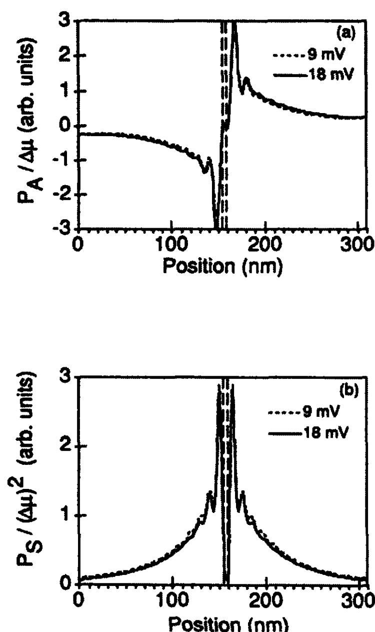
$$\mu_J(z) = \frac{\int E\, J(E,z)\, dE}{\int J(E,z)\, dE}. \tag{6.1}$$

Here the current density per unit energy is denoted by J(E,z). With such a definition, the local power transferred by the transport electrons to the lattice (denoted by P) can be calculated from the local current density times the change in μ_J along the direction of current:

$$P = -\int J(E,z)\, dE \cdot \frac{d\mu_J}{dz} = -J_{tot} \cdot \frac{d\mu_J}{dz}\ . \tag{6.2}$$

where z is the distance in the transport direction (Lake, Datta (1992), Eq. (10)). The Peltier heating/cooling is then found by taking the antisymmetric part of P (denoted by P_A), while the Joule heating is found from the symmetric part of P (denoted by P_S).

A theoretical model of the two effects (using a non-equilibrium Green's functions formalism with a local phonon model) for tunneling electron transport through a 5 nm barrier is shown in Figure 112. Here $P_A/\Delta\mu$ is the ratio of the antisymmetric (Peltier) part of the heating/cooling power to the applied chemical potential (voltage), and $P_S/\Delta\mu^2$ is a corresponding quantity for the symmetric (Joule) heating.



from "Energy balance and heat exchange in mesoscopic systems"
R. Lake, S. Datta, Phys. Rev. B, v46 p4757 (1992)

Figure 112. The Peltier heating/cooling and the Joule heating powers across a 5 nm tunneling barrier (Lake et al. 1992). Some of the transport energy of the hot electrons goes into the (top image, antisymmetric, reversible) Peltier effect, and some goes into Joule heating (bottom image, symmetric, irreversible). Two biases ($\Delta\mu$=9 mV, 18 mV) are checked.

In all approaches we have seen (19th century phenomenological thermodynamics, Onsager first order non-equilibrium thermodynamics with transport coefficients computed using the Boltzmann transport equation's assumptions, mesoscopic transport from non-equilibrium Green's function computations), the Peltier coefficient describes a thermodynamically reversible effect. For example, in Figure 112 the tunneling barrier seen by the current is has a reflection symmetry with respect to the center of the barrier, but the heating and cooling profile of the Peltier effect (upper image) do not. Thermodynamic reversibility requires time-reversal symmetry, which in this case requires that both the sign of the current (with units of

[charge/time]) and the Peltier power (which has units of [energy/time]) change together under time reversal. Therefore the Peltier coefficient defined as the antisymmetric part of the heating profile is consistent with the traditional understanding of the Peltier effect as thermodynamically reversible heating/cooling.

In the current theory on which the design (and expectations of enhanced ZT) of the metal superlattice ErAs system was based this does not seem to be the case[166]. The Peltier coefficient is evidently calculated as the difference between mean energies of the energies of the hot electrons emitted over the barrier and the equilibrium Fermi level in the emitter:

$$\pi = \mu_J - \mu \, . \tag{6.3}$$

The reversibility of the Peltier coefficient calculated in this manner seems questionable, since (as far as we can tell) no account has been taken of the heat and entropy that is produced when the electrons thermalize with the lattice on the collector side of the barrier. In other words, the Peltier coefficient calculated in this manner seems to include both the symmetric and anti-symmetric parts of $\frac{d\mu_J}{dz}$. Some of these considerations may be related to the conclusion of Ulrich et al. that for all known materials, thermoelectric refrigeration is more effective and efficient than single-barrier solid-state thermionic refrigeration[42]. It may be that the presence of multiple barriers is the necessary ingredient for effective thermionic effects predicted by Vashaee and Shakouri. We feel this point needs to be clarified, particularly with regards to accounting for the question of why (in the case of thermionic refrigeration) the extra heat produced when the hot electrons eventually come to thermal equilibrium with the lattice does not flow back to the emitter through the thermal conductivity of the lattice and electrons, canceling out some of the calculated Peltier heat flow. It would also be reassuring to see the validity of the Boltzmann transport assumption checked against a full quantum transport theory, since the Boltzmann equation does not always accurately describe electronic transport in superlattices[167]. An important step forward in this area was taken by Zebarjadi et al[168] using a Monte-Carlo analysis of thermoelectric transport perpendicular to thin-film heterostructures. Very little enhancement in the Seebeck coefficient due to thermionic barrier transport was reported (optimized ballistic Seebeck from model=315 μV/K, bulk Seebeck=295 μV/K). It is not clear to us how this is to be reconciled with the earlier theoretical predictions of greatly enhanced ZT, especially in view of the fact that the Monte Carlo results also seem to predict a steeply increasing electrical resistivity in the thermionic barrier regime.

In Chapter 5 we presented measurements made using the technique of thermoreflectance lock-in CCD imaging, a non-contact means of thermography capable of sub-micron resolution. We presented (and experimentally verified) a complete description of the measurement accuracy in terms of fundamental noise sources in the camera, and captured this model in a single analytic function that will hopefully be used by others in the field. We have demonstrated a thermal resolution of 18 mK and checked the thermoreflectance measurement against an independent measurement using a 3-ω setup built for that purpose. As a result of this work, we think we have advanced the ability of researchers to make quantitative measurements using this technique.

We have used micro/nano-scale thermoreflectance imaging for a series of measurements on polysilicon-gate FET devices and a MAM Bi_2Te_3 thermoelectric element imaged in the cross plane. The FET devices were used to investigate the general potential of the technique for the imaging of active semiconductor devices. The measurements provided an accurate means of understanding thermal phenomena in the devices. The saturation and linear regimes in an FET transistor could be easily distinguished, and the saturation heating in a 5 μm gate-length transistor was imaged with a resolution that saturated the capabilities of our optics. The possible applicability of the technique for reliability testing and for imaging dynamic heating in microprocessors were suggested with measurements of a point defect in the gate of a 5 μm FET and an image of the RC heating in the gate of an FET. We feel there is great potential applicability of this work towards device characterization and device reliability studies. We present two examples of important questions that might be addressed with this imaging technique. In Figure 113, a thermoreflectance image taken on pentacene-channel organic FET devices[169] is shown.

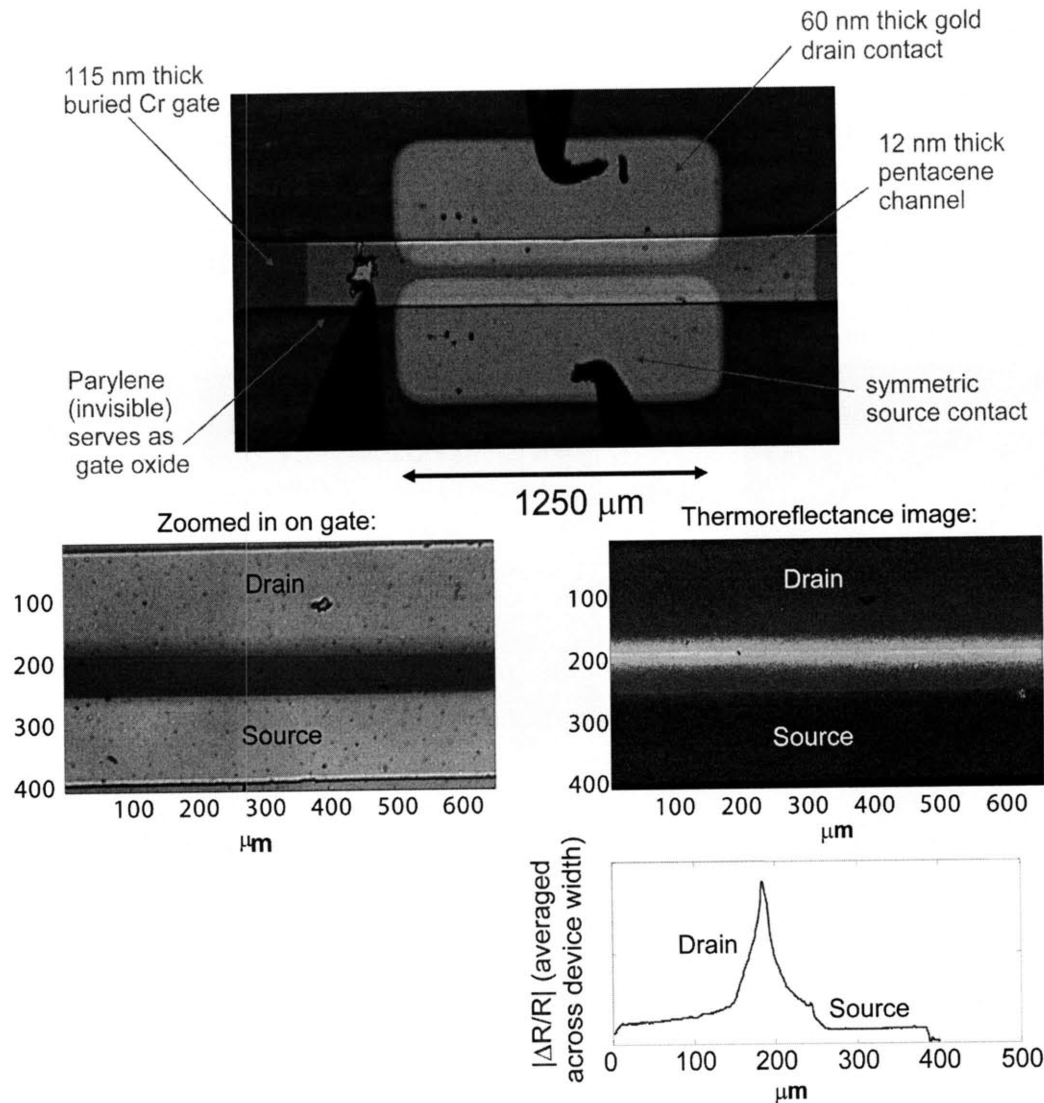


Figure 113. Image of contact heating in the drain of a pentacene channel organic FET device. (device courtesy of John Kymissis, Bulovic group, MIT)

One difficulty in the measurement of these devices is the extraction of electrical contact resistance from the IV characteristics of the devices. Typically devices of different channel lengths must be grown and the contact resistance extracted from a regression analysis on the IV characteristics of different length devices[170]. This is less than ideal, because the contact resistance between devices must be carefully controlled for this method to work. If the heating in the gold contacts could be directly measured, determination of the contact resistance from a single device might be possible. This would require both a calibration of the thermoreflectance and a model for the thermal impedance seen by the heat in the contact.

Another potential application for thermoreflectance imaging is the study of degradation mechanisms in HEMTs (High Electron Mobility Transistors)[171-173]. A high resolution (Sparrow criteria resolution = 316 nm) thermoreflectance image of such a device under bias is shown in Figure 114.

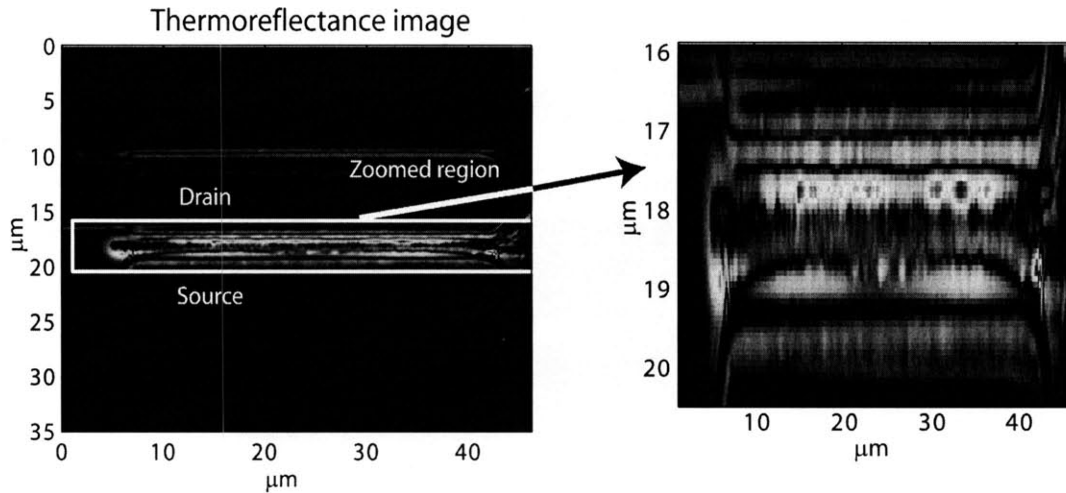


Figure 114. Heating in a GaAs HEMT. (device courtesy of Anita Villanueva, del Alamo group, MIT)

The exact causes of the degradation of these devices is an area of on-going research, but long-term reliability of these devices is related to a number of effects. Two of these mechanisms that might be well suited for study using thermoreflectance imaging are hot electron degradation in the drain[173] and piezoelectric stress build-up (due to a chemical degradation process) under the gate[171]. The thermoreflectance measurement might be sensitive to both a bias-dependent piezoelectric strain and also to normal thermal effects in these devices, since both could conceivably affect the dielectric response of the material. The complexity of both the multilayer HEMT structure and the possible interplay of differing thermoreflectance mechanisms make the measurement challenging, but potentially very interesting. The extremely small size of the devices might make the development of a deep blue or UV-based thermoreflectance system (with a back-thinned CCD array in order to achieve short wavelength sensitivity) a worthwhile goal.

The measurement of the MAM Bi_2Te_3 demonstrated the first thermoreflectance imaging a thermoelectric element in the cross-plane direction. While quantitative measurements of the

temperature were possible, quantitative modeling of the expected temperature in the device using a finite-difference time domain solver was complicated by the sensitivity of the measurement to thermal boundary conditions whose magnitude could only be estimated, and to the multicrystalline character of the sample. We judged that the match between experiment and theory in this case was not yet good enough for quantitative thermoelectric property characterization. However, if the heat transfer at the contacts can be understood, the thermoelectric properties of thin films might be extracted from the measured temperature distributions. We explored this possibility using the ErAs nanodot thin-film sample depicted in Figure 115.

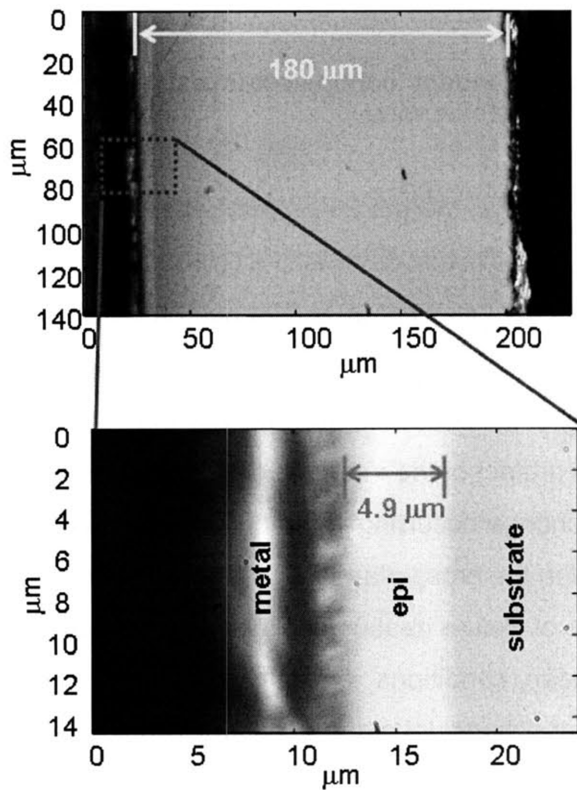


Figure 115. Photomicrograph of the cross-plane view of an ErAs nanodot sample. The 4.9 μm superlattice is grown atop a 175 μm InP substrate.

The results of a thermoreflectance measurement on this structure are shown below in Figure 116.

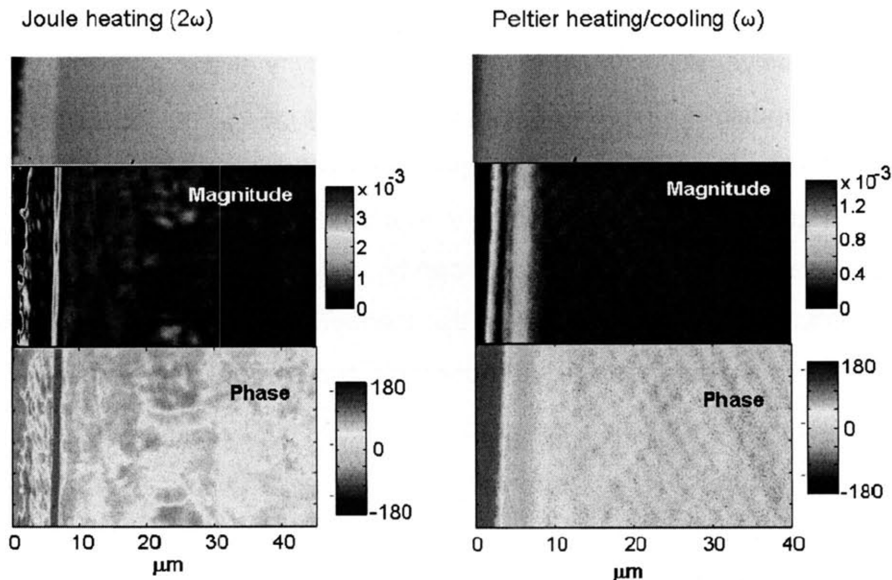


Figure 116. Cross-plane thermoreflectance images of ErAs nanodot thermoelectric material on InP substrate. The current excitation was a 5 A amplitude, 5.5 Hz sine wave.

The results shown in Figure 116 are interesting, but are also somewhat ambiguous. A sharp feature in the Joule heating profile is detected between the thin-film and the substrate. This might be due to real heating at the interface, but owing to its small size, we feel more tests are needed to determine whether the heating is real or an artifact of the interface due to sample motion. Likewise, there is a clear 180 degree phase change occurring in the phase image of the Peltier heating, similar to that which was observed in the cross-plane Bi_2Te_3 MAM images of Chapter 5. It is quite possible that this is the signature of Peltier heating/cooling, but owing to the uncertainty in the thermal contact resistance boundary conditions, it is difficult to correlate the position of this phase switch with our theoretical predictions, let alone use it to extract information about the material parameters. Also unexplained in this image are the strange features visible in the substrates of all of the images; in the joule heating images, the substrate heating appears blotchy, and in the Peltier heating images there is a regular, tiled appearance to the thermal response.

We also imaged the device on a larger scale (including the entire substrate) in order to attempt to match the theoretically expected heating profiles (using the thin-film and substrate material parameters assumed in Chapter 4.6 and the cross-plane time-domain finite difference

simulation routine introduced in Chapter 5.8) with the cross plane thermoreflectance magnitude and phase. The results are shown in Figure 117.

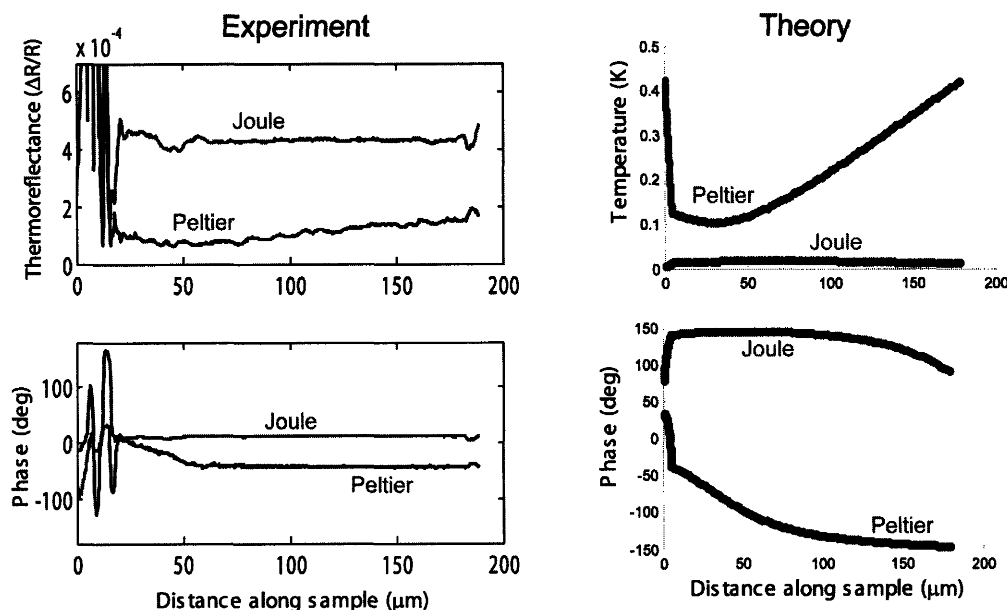


Figure 117. Cross-plane thermoreflectance on a 4.9 micron ErAs nanodot thin-film on a 180 micron InP substrate. 8 A pk-pk current producing a heating frequency of 3 Hz was was used to generate the temperature profiles.

While there is some qualitative agreement, especially between the Peltier phase and magnitude profiles and their simulated counterparts, the presence of the artifacts discussed earlier and the unexpectedly large Joule effects relative to Peltier effects suggest that further work is needed before the technique is ready for thermoelectric material parameter extraction. The combination of the high resolution needed to resolve the 4.9 μm thin-film and the long measurement times needed to accurately measure the small thermal signals (100 mK range) make the measurement rather more challenging than any of the other microelectronic devices we have examined. We would suggest that future work in this area be directed towards the cross-plane thermoreflectance imaging (and parameter extraction) of a bulk resistive material of known resistivity of at least 100 μm in length. The theory we have developed can be used to match the expected Joule heating response very accurately. Using the known resistivity, the values for the thermal resistances at the contacts (important for quantitative measurements) can be studied, and hopefully minimized through a careful redesign of the sample holder. If confidence in the size and repeatability of the thermal contacts, the next step would be to image a bulk

thermoelectric material of at least 100 μm in length. The key properties for successful measurement are that the material be a reasonably good thermoelectric material (so that Peltier heating and cooling are observable) and that a good polished surface or cleaved facet can be obtained. An optimally doped ternary or quaternary III-V semiconductor like InGaAs might be a good candidate because its thermal resistance is relatively high, and a reasonable cleave can usually be obtained. A single crystal Bridgeman grown sample from the $(Bi_{1-x}Se_x)_2Te_3$ system might also be a good choice. When these measurements can be performed successfully and are found to be in good agreement with the theory we have developed, then we think the measurements on thin-films could be performed with the reasonable expectation of quantitative (< 10% error) parameter extraction.

With this thesis, we've tried to make some contributions to the fields of thermoelectric device design, thermoelectric property characterization, and micro/nano-scale thermal imaging. The field has been an exciting one to be a part of, with new ideas from materials science and solid state transport continually being used to motivate new materials and growth techniques, which are in turn measured with new material and device characterization techniques. We think that with proper attention to system level design constraints, useful new engineering solutions are already possible with current TE materials. There has been a great deal of progress in the last several years in new material development, particularly towards the engineering of phonon scattering nanostructures to drive down material thermal conductivity while maintaining the thermoelectric power factor ($\alpha^2\sigma$). The future prospects for thermoelectric technology seem bright.

Appendix 1: Justification of the standard linear treatment of the quantization process

As mentioned earlier, quantization is an inherently nonlinear process. The signal flow diagram and transfer function for the process are shown in Figure 118. A signal to be measured (x) is combined with some additive noise (v), and the resultant signal (w) is put through a quantizer. The quantized output (y) is related to the original signal by an error term (ε) which has error due to the additive noise and the quantization process.

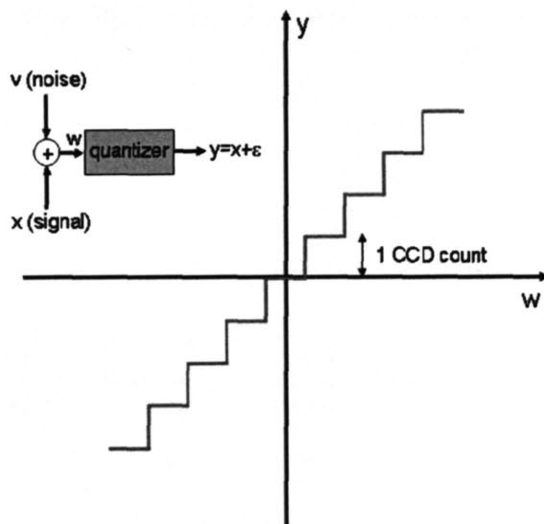


Figure 118. Quantization process. The quantizer effectively rounds off the input signal w to the nearest quantization step.

For what follows, we assume that the noise signal v is Gaussian, white, and has a standard deviation of σ_d. In other words, v=d from the discussion of thermoreflectance.

We would like to be able to treat the error ε like a simple independent noise source similar to v. Unfortunately, the error ε is generally not independent of the signal x. In fact, if w is so small that it never crosses into any adjacent quantization levels, the signal and error are exactly negatively correlated (ε=-x). In such a situation, assuming that ε could be modeled as a zero mean Gaussian noise source would be completely wrong. However, if the standard deviation of v is much greater than the quantization bin size, we know that the statistics of ε will begin to

approach those of v. A quantitative measure of when this change occurs can be made by calculating the correlation of x and ε, given by $E[x\varepsilon]$, where E denotes the expectation value.

From Figure 118, we can write down an expression for the pdf of ε conditioned on a known value of the signal[153]

$$p_{\varepsilon|x} = \sum_{k=-\infty}^{\infty} \delta(\varepsilon + x - k) \int_{-1/2+k}^{1/2+k} p_v(w - x)dw . \tag{A1.1}$$

Here p_v denotes the pdf of the noise v. The probability distribution takes the form of a sum of delta functions at discrete values of k corresponding to each of the output quantization levels of the A-to-D converter. The number of quantization levels is assumed infinite so that each bin can be presumed identical. This should be fine as long as the signal does not clip (we stay well clear of the maximum and minimum quantization levels). The integral in the expression calculates the weight (likelihood) of each possible output quantization bin according to the probability that the sum of the input and noise falls in the range corresponding to that bin. For Gaussian input noise, the integrals can be analytically expressed in terms of the error function.

Substituting in the Gaussian pdf and doing the integral gives

$$p_{\varepsilon|x} = \frac{1}{2}\sum_{k=-\infty}^{\infty} \delta(\varepsilon + x - k)\left\{ \mathrm{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[\frac{1}{2} + k - x\right]\right) - \mathrm{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[-\frac{1}{2} + k - x\right]\right)\right\}. \tag{A1.2}$$

Once this is known, the correlation between ε and x can be found according to

$$E[x\varepsilon] = \int_{-\infty}^{\infty} x\varepsilon\left(p_{\varepsilon|x}\right)d\varepsilon . \tag{A1.3}$$

The properties of the delta function are used to write the result in terms of a sum over all of the quantization levels.

$$E[x\varepsilon] = \frac{x}{2}\sum_{k=-\infty}^{\infty} (k - x)\left\{ \mathrm{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[\frac{1}{2} + k - x\right]\right) - \mathrm{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[-\frac{1}{2} + k - x\right]\right)\right\} \tag{A1.4}$$

This sum can be explicitly computed, and the results are summarized in Figure 119.

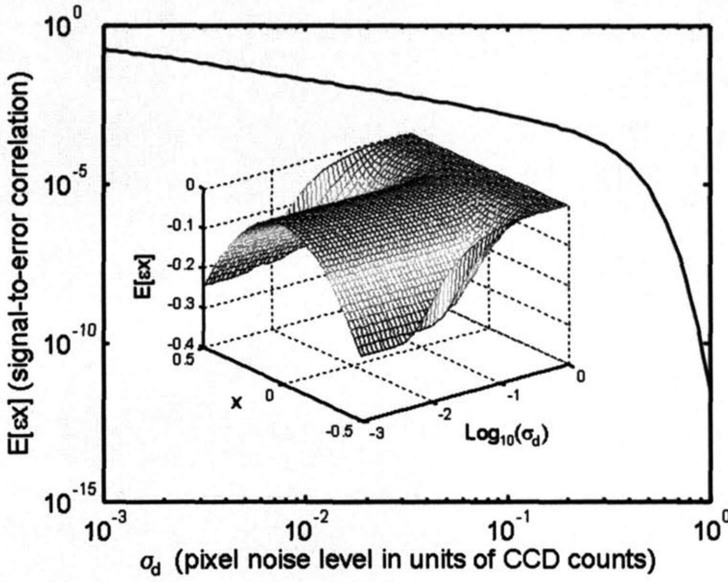


Figure 119. The worst case dependence of the correlation between the quantization error and the signal on the input noise level. The inset shows the correlation as a function of the input noise level and the position of the signal in the quantization bin (x).

The correlation everywhere is negative, as we expect. The result of this analysis is that for noise levels on the order of the quantization bin size or larger, the correlation between the error and the input becomes incredibly small (note the logarithmic scale), justifying the treatment of ε as an independent noise source. Above this threshold, the variance of ε is well approximated by the standard formula for quantization noise

$$\sigma_\varepsilon^{\,2} = E[\varepsilon^2] - E[\varepsilon]^2 \cong \sigma_d^{\,2} + \frac{1}{12}. \tag{A1.5}$$

This can be confirmed by plotting the first and second moments of ε using summations similar to Eq. A1.4:

$$E[\varepsilon] = \frac{1}{2}\sum_{k=-\infty}^{\infty}(k-x)\left\{\operatorname{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[\frac{1}{2}+k-x\right]\right) - \operatorname{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[-\frac{1}{2}+k-x\right]\right)\right\} = 0 \tag{A1.6}$$

$$E[\varepsilon^2] = \frac{1}{2}\sum_{k=-\infty}^{\infty}(k-x)^2\left\{\operatorname{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[\frac{1}{2}+k-x\right]\right) - \operatorname{erf}\left(\frac{1}{\sigma_d\sqrt{2}}\left[-\frac{1}{2}+k-x\right]\right)\right\} \tag{A1.7}$$

The true noise variance σ_ε^2 is shown in , and compared with the standard approximation of Eq. A1.5. The formula is very accurate for input noise signals greater than around half the quantization bin size.

Appendix 2: Thermoreflectance phase noise analysis

From the standpoint of temperature measurement, knowledge of the pdf of the measured thermoreflectance signal is sufficient for most practical purposes, but for more general applications of the thermoreflectance technique, information on the phase distribution may be useful. The phase of the thermoreflectance is a useful quantity for time domain measurements (e.g. imaging thermal waves[123]) and for checking that a measurement has converged on a region known to be at a uniform temperature[143].

It is convenient to define new random variables according to

$$\begin{aligned} Y &\equiv A - B, \quad & p_Y &= \mathcal{N}(\mu_A - \mu_B, 2\sigma^2) = \mathcal{N}(\mu_Y, \sigma_\phi^2) \\ X &\equiv A + B \quad & p_X &= \mathcal{N}(\mu_A + \mu_B, 2\sigma^2) = \mathcal{N}(\mu_X, \sigma_\phi^2) \end{aligned} \tag{A2.1}$$

such that

$$\phi = \arctan\left(\frac{Y}{X}\right) - \psi = \arctan\left(\frac{I_1 - I_2 - I_3 + I_4}{I_1 + I_2 - I_3 - I_4}\right) - \psi. \tag{A2.2}$$

Y and X so chosen can be shown to be independent by writing the joint pdf for A and B, substituting variables, and then factoring $p_{X,Y}(X,Y) = [p_X(X)][p_Y(Y)]$.

By writing the phase pdf as in Eq. A2.2, it is recognizable as the angle formed by the vectorial sum of the Gaussian rvs X and Y, which can be obtained by integrating Eq. 5.4.17 over r, where now $r \equiv \sqrt{X^2 + Y^2}$. (Similarly, the angle θ in Eq. 5.4.17 can be identified with the phase ϕ.) One finds:

$$p_\phi = \frac{1}{2\pi}\exp\left(-\frac{\mu_X^2 + \mu_Y^2}{2\sigma_\phi^2}\right) + \tag{A2.3}$$

$$\frac{\sqrt{\mu_X^2 + \mu_Y^2}\cos(\phi - \theta)}{2\sqrt{2\pi}\sigma_\phi}\exp\left(-\frac{(\mu_X^2 + \mu_Y^2)\sin^2(\phi - \theta)}{2\sigma_\phi^2}\right)\left(1 + \operatorname{erf}\left(\frac{\sqrt{\mu_X^2 + \mu_Y^2}\cos(\phi - \theta)}{\sqrt{2}\sigma_\phi}\right)\right) \text{Here}$$

$$\tan\theta \equiv \frac{\mu_Y}{\mu_X}.$$

The pdf has the form of a uniform term which exponentially decreases with increasing signal to noise level, and a term peaked around the angle θ. When the noise σ_ϕ is allowed to approach

zero, the pdf approaches zero everywhere except around $\phi - \theta = 0$, where it diverges. Thus the phase pdf converges to the value expected for the case of no noise (Eq. 5.4.4). Figure 120 shows the effects of removing noise on a phase pdf centered on θ=0.

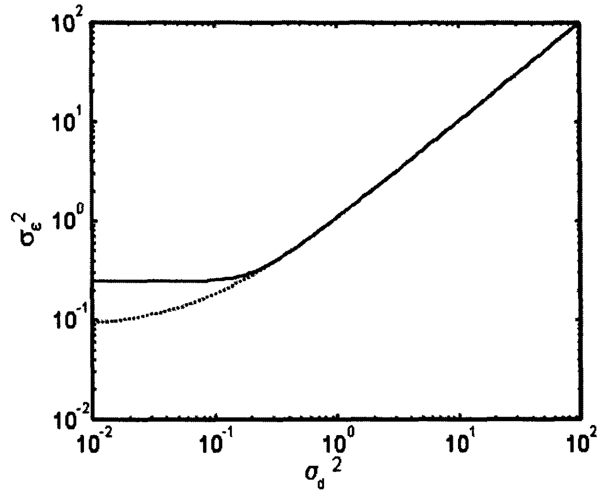


Figure 120. The second moment of the quantization error as a function of the input noise level (solid line). Also shown is the moment computed using the standard approximation σ_d^2 +1/12 (dotted line).

Appendix 3: Marlow MAM characterization data

From Marlow Industries:

SINGLE TE PELLET — DATE: 11/12/2004

PELLET WIDTH 1:	0.15
PELLET LENGTH:	0.19
PELLET WIDTH 2:	0.15

TEMP K	ALPHA (V/DEG C)	KAPPA (W/CM/DEG C)	RHO (OHM * CM)	Z (/DEG K)
337.36	-2.4194E-04	1.5126E-02	1.5223E-03	2.5421E-03
320.88	-2.4226E-04	1.4841E-02	1.4324E-03	2.7608E-03
308.06	-2.4029E-04	1.4688E-02	1.3567E-03	2.8975E-03
292.93	-2.3756E-04	1.4786E-02	1.2674E-03	3.0114E-03
279.09	-2.3387E-04	1.5046E-02	1.1878E-03	3.0605E-05
262.24	-2.2664E-04	1.5237E-02	1.0887E-03	3.0964E-03
247.94	-2.2256E-04	1.5919E-02	1.0057E-03	3.0940E-03
232.43	-2.1372E-04	1.6267E-02	9.1975E-04	3.0530E-03
215.62	-2.0744E-04	1.7481E-02	8.2820E-04	2.9723E-03
197.56	-1.9603E-04	1.8266E-02	7.3226E-04	2.8732E-03
185.85	-1.9336E-04	1.9980E-02	6.7288E-04	2.7811E-03
178.89	-1.8580E-04	2.0066E-02	6.3826E-04	2.6955E-03
300.00	-2.3798E-04	1.4726E-02	1.3059E-03	2.9451E-03
256.00	-2.2517E-04	1.5493E-02	1.0523E-03	3.1097E-03

DATA RECORDED IN FILE C:\ZTestWin\Data\T111204A.dat

This characterization data was provided for a large sample of Marlow's material. The small elements characterized in this work were not previously measured, but are supposed to be made from identical material.

Appendix 4: Thermal quadrupoles[55] used to estimate to meter bar errors

Suppose that we terminate a 5 cm long copper cylinder (area=A, thermal conductivity=K=4.0 W/cm/K, thermal diffusivity=a=1.1 cm^2/s) on one end with a heat sunk boundary condition consistent with forced convection with chilled water (h=0.6 $W/cm^2/K$). This is the geometry of the meter bar used for the single element heat flux measurements in Chapter 4. An important issue for the accuracy of the measurement is the time is takes to establish steady-state temperature conditions in the bar. In other words, given a sudden change in temperature at the top of the bar, how long must one wait until the heat flux in the bar can be measured to an accuracy of 10%? The question can be answered using the thermal quadrupole method.

Consider a uniform cylindrical slab of material, of length L, area A, and thermal diffusivity a (Figure 121). Suppose we would like to find the thermal impedance of the cylinder as a function of thermal excitation frequency for an arbitrary thermal boundary condition on one end. We can begin by finding the transient thermal quadrupole characterizing the slab under conditions of uniform axial heat flux. We denote the temperature and heat in the cylinder as T(x,t) and Q(x,t). The temperature and heat at the input face (x=0) we denote by T(x=0,t) and Q(x=0,t), while on the output face (x=L) we have T(x=L,t) and Q(x=L,t).

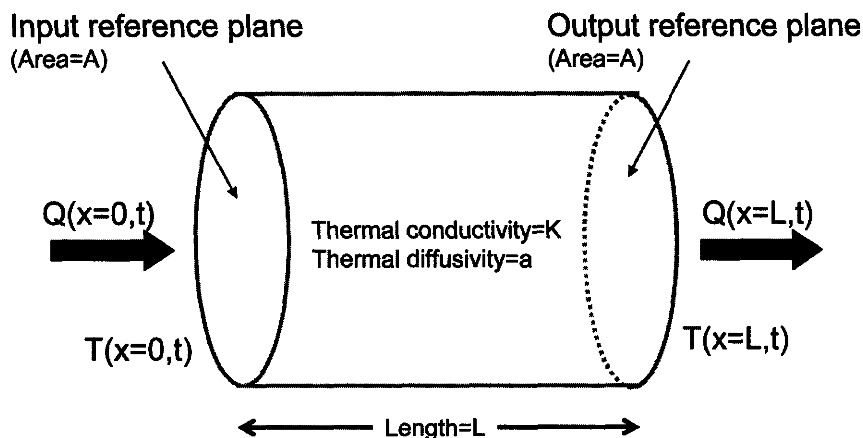


Figure 121. Cylindrical geometry considered for time dependent one dimensional heat transfer.

Defining $T(x,s)=\int_0^\infty T(x,t)e^{-st}dt$ and $P(x,s)=\int_0^\infty Q(x,t)e^{-st}dt$, we may write the desired input-output relation as

$$\begin{bmatrix} T(x=L,s) \\ P(x=L,s) \end{bmatrix} = M(s) \begin{bmatrix} T(x=0,s) \\ P(x=0,s) \end{bmatrix}. \quad \text{(A4.1)}$$

We are interested in solving for the elements of the quadrupole matrix

$M(s) = \begin{bmatrix} m_{11}(s) & m_{12}(s) \\ m_{21}(s) & m_{22}(s) \end{bmatrix}$, relating the transforms of the input and output temperatures and heat fluxes in complex frequency space.

The time dependent heat equation in one dimension is given by

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{a}\frac{\partial T}{\partial t} . \quad \text{(A4.2)}$$

The heat passing through any circular cross section of the cylinder is given by

$$Q = -KA\frac{\partial T}{\partial x}. \quad \text{(A4.3)}$$

We can find the Laplace transform of both of these equations:

$$\frac{\partial^2 T}{\partial x^2} = \frac{s}{a}, \quad P = -KA\frac{\partial T}{\partial x}. \quad \text{(A4.4)}$$

The general solution of the equation for transformed temperature □ can then be written in terms of exponentials (or sinh and cosh functions) as

$$T(x,s) = C_1 \sinh(\sqrt{s/a}\, x) + C_2 \cosh(\sqrt{s/a}\, x). \quad \text{(A4.5)}$$

The corresponding solution for the transformed heat flux can then be found:

$$P(x,s) = -KA\sqrt{s/a}\left(C_1 \cosh(\sqrt{s/a}\, x) + C_2 \sinh(\sqrt{s/a}\, x)\right). \quad \text{(A4.6)}$$

By treating the values of the temperature and heat flux at x=0 as known, we can solve for the two constants C_1 and C_2 that are needed to calculate the temperature and heat flux at x=L:

$$C_2 = T(x=0,s)$$
$$C_1 = \frac{-P(x=0,s)}{KA\sqrt{s/a}}. \quad \text{(A4.7)}$$

Calculating the matrix M evaluated at x=L is straightforward:

$$M(s) = \begin{bmatrix} m_{11}(s) & m_{12}(s) \\ m_{21}(s) & m_{22}(s) \end{bmatrix} = \begin{bmatrix} \cosh\left(L\sqrt{s/a}\right) & \dfrac{-\sinh\left(L\sqrt{s/a}\right)}{KA\sqrt{s/a}} \\ -KA\sqrt{s/a}\sinh\left(L\sqrt{s/a}\right) & \cosh\left(L\sqrt{s/a}\right) \end{bmatrix}. \quad \text{(A4.8)}$$

Knowledge of the matrix M is sufficient for many calculations, but it is instructive and useful to relate the matrix form of the solution to a two-port impedance model familiar from electrical circuit theory.

It is straightforward to show that the because the thermal system is passive, the determinant of the matrix M is unity[55]. Because of this constraint, the four element matrix can be represented in an equivalent circuit model with three impedances, connected in a "T" as shown in Figure 122.

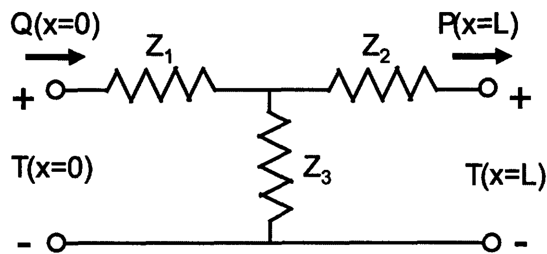


Figure 122. Impedance "T" description of passive thermal two-port.

If the impedances Z1, Z2, and Z3 are taken as

$$Z_1 = \frac{1-m_{22}}{m_{21}}, \quad Z_2 = \frac{1-m_{11}}{m_{21}}, \quad Z_3 = -\frac{1}{m_{21}}, \tag{A4.9}$$

then one can show that

$$m_{11} = 1+\frac{Z_2}{Z_3}, \quad m_{12} = -Z_1 - Z_2 - \frac{Z_1 Z_2}{Z_3}, \quad m_{21} = -\frac{1}{Z_3}, \quad m_{22} = 1+\frac{Z_1}{Z_3}. \tag{A4.10}$$

With this choice, the constraint det|M|=1 is automatically satisfied, leaving three of the four matrix elements linearly independent of one another. For the particular case (uniform 1-D transport) solved above, we find that

$$Z_1 = Z_2 = \frac{\cosh\left(L\sqrt{s/a}\right)-1}{KA\sqrt{s/a}\,\sinh\left(L\sqrt{s/a}\right)}, \quad Z_3 = \frac{1}{KA\sqrt{s/a}\,\sinh\left(L\sqrt{s/a}\right)}. \tag{A4.11}$$

For a general passive two port, Z_1, Z_2, and Z_3 will be different. The fact that Z_1=Z_2 here can be shown to be a consequence of the symmetry of the problem; the cylinder looks the same from the input or the output, and the circuit model reflects that symmetry.

To model the meter bar system, we will terminate the thermal quadrupole with a heat sink whose thermal resistance is 1/(hA). This can be represented as a thermal resistance "to ground" as shown in Figure 123.

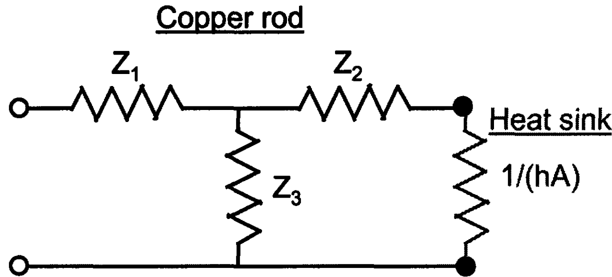


Figure 123. Impedance model of copper rod terminated by a heat sink.

Calculating the thermal impedance of the network at the input port gives

$$Z_{th} = Z_1 + Z_3 \,||\, \left(Z_2 + \frac{1}{hA} \right) = \frac{\dfrac{1}{KA\sqrt{s/a}} \sinh\left(L\sqrt{s/a}\right) + \dfrac{1}{hA} \cosh\left(L\sqrt{s/a}\right)}{\dfrac{KA\sqrt{s/a}}{hA} \sinh\left(L\sqrt{s/a}\right) + \cosh\left(L\sqrt{s/a}\right)} . \qquad \text{(A4.12)}$$

The steady-state value of the thermal impedance is given by the limit of $Z_{th}(s)$ as $s \rightarrow 0$. Using L'Hopital's rule, we can confirm

$$Z_{th}\left(s \to 0\right) = \lim_{s \to 0} \left(\frac{\sinh\left(L\sqrt{s/a}\right)}{KA\sqrt{s/a}} \right) + \frac{1}{hA} = \frac{L}{KA} + \frac{1}{hA} . \qquad \text{(A4.13)}$$

This is consistent with our usual intuition for the steady-state. The absolute value of the complex impedance can be plotted as a function of frequency by setting $s=j\omega=j2\pi f$ in Eq. A4.12. The results are shown below in Figure 124.

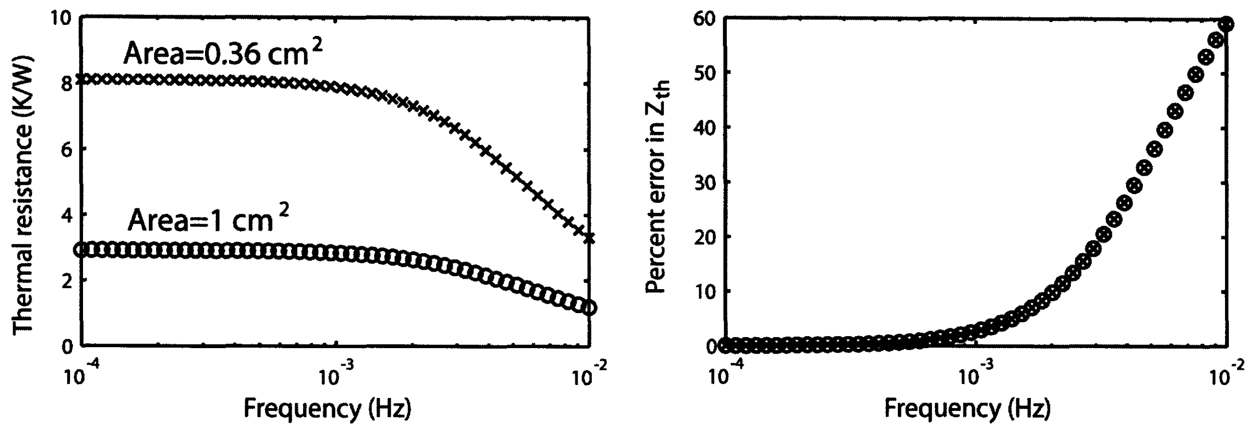


Figure 124. Thermal impedance of two heat-sunk copper bars of differing cross-sectional area, plotted as a function of frequency. The percent difference between each thermal impedance and its steady-state value is also plotted. The analysis is useful for determining the amount of time needed to perform meter bar heat flux measurements.

The results for a cross sectional area of A=0.36 cm^2 are shown with the 'x' symbol, and those for an area of 1 cm^2 are shown with the circles. It is evident from the leftmost plot in Figure 124 that at low enough frequencies both thermal impedances settle to their steady-state values, and at high frequencies the thermal resistances fall. If the temperatures at the ends of the meter bar are used to make a measurement of the steady-state heat flux, the time dependence of the thermal impedance will cause an error unless enough time is left for the bar's temperature to settle. In the rightmost plot in Figure 124 the percent difference of the thermal impedances from their steady-state values is plotted versus frequency. Above around 2 mHz (i.e. at time scales less than around 500 seconds), the errors in measured heat flux will be more than 10%. It is also clear that the error is independent of the meter bar area.

Appendix 5: Details of microchannel flow (calculation of Nu and f_D)

To calculate the pressure drop (and pumping power) in the microchannel, the Darcy friction factor f_D must be calculated. Similarly, to calculate the heat transfer coefficient at the wall of the channel, the mean Nusselt number Nu must be calculated in the channel. There are different correlations available (empirical formulae) for both dimensionless numbers that are valid in different regimes (e.g. laminar vs. turbulent). We have built a model of the channel flow by patching together the empirical correlations published for each regime. The book by Incropera[48] is an excellent resource, and we rely on it for all of the flow correlations used here. It is worth pointing out that due to their empirical nature and due to the complex nature of water flow, these correlations should always be taken with a grain of salt. For example, we assume that the transition between laminar and turbulent flow occurs at a Reynolds number of Re=2300, consistent with Incropera's value for macroscopic pipe flow (Incropera, p70). Some authors have seen the onset of turbulence at lower apparent Reynolds numbers in small microchannels, while some have not. It is now widely believed that these differences are due to subtle entrance effects and channel wall roughness effects, rather than any fundamentally new physics. Obviously, the choice of how to model the onset of turbulence is an important one. For a real system, it would be best to measure the Reynolds number corresponding to the minimum turbulent flow velocity and adjust the model accordingly. We feel that at the present time we have done the best we can with published correlations.

Following the calculations in Chapter 3 of this thesis, we assume a channel height b, a channel width a, a channel length L, and a mean flow velocity u. We define a distance measured from the start of the channel in the direction of channel flow as x. The following definitions are also useful:

$$\text{Hydraulic diameter} \equiv D_h = \frac{4 \cdot \text{channel area}}{\text{channel perimeter}} = \frac{4ab}{2a+2b} \qquad \text{(A5.1)}$$

$$\text{Aspect ratio} \equiv A_R = \min\{a/b, b/a\} \qquad \text{(A5.2)}$$

<u>Friction factors</u>

As the water flows down the channel, the flow velocity profile gradually changes from uniform to parabolic as fully developed hydrodynamic flow conditions are approached. This change happens over the course of a distance known as the hydrodynamic entrance length, which we call x_{fd}. For pipe flow, the hydrodynamic entrance length is typically around (Incropera, p72):

$$x_{fd} = 0.08 \cdot D_h \cdot Re. \qquad (A5.3)$$

In light of this observed dependence, it is useful to define a scaled hydrodynamic distance down the channel according to:

$$x^+ = x/(D_h \cdot Re). \qquad (A5.4)$$

In Incropera (page 78) one finds the following table for apparent mean friction factors for developing laminar flow in rectangular channels of varying aspect ratios:

x^+	f_D*Re/4			
	A_R=1.0	A_R=0.5	A_R=0.2	A_R<=0.1
0.001	111.0	111.0	111.0	112.0
0.003	66.0	66.0	66.1	67.5
0.005	51.8	51.8	52.5	53.0
0.007	44.6	44.6	45.3	46.2
0.009	39.9	40.0	40.6	42.1
0.010	38.0	38.2	38.9	40.4
0.015	32.1	32.5	33.3	35.6
0.020	28.6	29.1	30.2	32.4
0.030	24.6	25.3	26.7	29.7
0.040	22.4	23.2	24.9	28.2
0.050	21.0	21.8	23.7	27.4
0.060	20.0	20.8	22.9	26.8
0.070	19.3	20.1	22.4	26.4
0.080	18.7	19.6	22.0	26.1
0.090	18.2	19.1	21.7	25.8
0.100	17.8	18.8	21.4	25.6
0.200	15.8	17.0	20.1	24.7
1.000	14.2	15.5	19.1	24.0

For example, at a fraction $x = 0.03 \cdot D_h \cdot Re$ into the entrance region of a square cross section (A_R=1) microchannel, the table tells us that the apparent Darcy friction factor can be found according to $f_D\left(x^+ = 0.030\right) = 24.6 \cdot 4/Re$. For x>1 (well into the fully developed flow region of the channel), one may use the value at $x^+ = 1$.

We have written code that does interpolated table lookup to these values, so that given a distance down the channel x, a Reynolds number Re, and a channel aspect ratio A_R, the table lookup function returns $f_D(x, Re, A_R)$. To account for the variation in the apparent friction factor

along the channel, the f_D at thirty equally spaced x-points along the channel are found using the table lookup function, and the average is taken.

There is a complication to the above system of calculating f_D due to the onset of turbulence. If the Reynolds number is greater than 2300, a different calculation for the friction factor is performed. The correlation used to model this effect is (Incropera, p79):

$$f_{D,turb}(x,D_h,Re,A_R) = 4 \cdot A(x)\left(Re\left[\frac{2}{3}+\frac{11}{24}A_R(2-A_R)\right]\right)^{B(x)}$$

$$A(x) = 0.0929 + \frac{1.0161}{x/D_h} \tag{A5.5}$$

$$B(x) = -0.2680 + \frac{0.3193}{x/D_h}$$

The same channel length averaging procedure can be used in this case. We have modified the turbulent correlation from Incropera in one significant way: we have added a small offset to the turbulent friction factor to make the transition between laminar and turbulent flow smooth.

Nusselt number

The other key fluid parameter for a microchannel cooler is the Nusselt number Nu of the flow. Similar to the fully developed flow distance, we can define a dimensionless fully thermal distance x^*:

$$x^* = x/(D_h \cdot Re \cdot Pr). \tag{A5.6}$$

In the laminar regime, we proceed in a similar manner to the friction factor calculation, using a table lookup with the aspect ratio and the distance into the microchannel (Incropera, p82).

x^*	Nu A_R=1.0	A_R=0.5	A_R=.33	A_R=.25	A_R<=0.1
0.00010	25.2	23.7	27.0	26.7	31.4
0.00250	8.9	9.2	9.9	10.4	11.9
0.00500	7.10	7.46	8.02	8.44	10.0
0.00556	6.86	7.23	7.76	8.18	9.8
0.00625	6.60	6.96	7.50	7.92	9.5
0.00714	6.32	6.68	7.22	7.63	9.3
0.00833	6.02	6.37	6.92	7.32	9.1
0.01000	5.69	6.05	6.57	7.00	8.8
0.01250	5.33	5.70	6.21	6.63	8.6
0.01670	4.91	5.28	5.82	6.26	8.5
0.02500	4.45	4.84	5.39	5.87	8.4
0.03300	4.18	4.61	5.17	5.77	8.3
0.05000	3.91	4.38	5.00	5.62	8.25
0.10000	3.71	4.22	4.85	5.45	8.24
1.00000	3.60	4.11	4.77	5.35	8.23

The post-turbulence Nusselt correlation is the modified Gnielinski correlation given in Incropera (p83).

$$Nu(x,f_D,Re,Pr,D_h)=\frac{(f_D/8)(Re-1000)Pr}{1+12.7\sqrt{f_D/8}\left(Pr^{2/3}-1\right)}\left[1+\left(\frac{D_h}{x}\right)^{2/3}\right]. \quad \text{(A5.7)}$$

Once again, we use the appropriate correlation at each point in the channel and use the length average in the system calculations. Again, the post-turbulence and pre-turbulence Nusselt numbers are forced to match up by adding a small offset to the post-turbulence Nusselt number.

Completing the model

Once the length averaged channel Nusselt number is calculated, the mean channel wall heat transfer coefficient is given by:

$$h=\frac{Nu\cdot\kappa}{D_h}, \quad \text{(A5.8)}$$

where κ is the thermal conductivity of the water.

The pressure drop is calculated from the Darcy friction factor according to:

$$\Delta P = \frac{1}{2}\rho u^2\left(2K_{90} + K_c + K_e + \frac{f_D \cdot L}{D_h}\right). \tag{A5.9}$$

$$K_{90} = 1.2, \quad K_c = 0.75, \quad K_e = 0.75$$

Here ΔP the pressure drop across a channel [Pa], ρ is the density of water [kg/cm^3], and K90, Kc, and Ke are dimensionless loss coefficients modeling two 90 degree bends, the entrance, and the exit, respectively (Incropera, p76).

Appendix 6: MATLAB code

This code contains a few of the more interesting or useful routines written to perform modeling for the work in this thesis. It is sorted by Chapter. It is also documented. Mostly. ☺

****

Chapter 2 code:

Finite difference solver for Onsager relations

```
function [Powerout, efficiency, R_int, Powerout_E, efficiency_E,
Z_est]=fun_fdtehs(Thpassed, Tfluidpassed, Lpassed, Hpassed,
Rconpassed, Rloadpassed, factors)
% this code is a 1-D finite difference solver for functionally graded
thermoelectrics
% implements a heat sink on the cold side, set H-> large for "ideal
temperature" boundary on cold side
%
% PARAMETERS
% Thpassed: hot side temperature (deg K)
% Tfluidpassed: cold side fluid temperature (deg K)
% Hpassed: heat sink thermal resistance (W/cm^2/K)
% Lpassed: length of the device (cm)
% Rconpassed: contact resistivity to the unicouple material from the
metal contact (ohm cm^2)
% Rload: this is the load resistance used to extract power; leave it
out or make it 'matched' for a perfectly matched load
% factors: scaling for [K alpha sigma]

% set to unity if none were passed
if nargin<7
   factors=[1 1 1];
end

% get load resistance, or assume matched if none was passed
matched_load=0;
if (nargin<6) | (Rloadpassed=='matched')
   matched_load=1;
    Rloadpassed='matched';
end
```

```
% set contact resistance to zero if none was passed
if nargin<5
   Rconpassed=1e-20;
end

% set contact resistance to zero if none was passed
if nargin<4
   Hpassed=1e8;
end

more off;
global A Rload Rcontacts Tfluid H
H=Hpassed;
Tfluid=Tfluidpassed;
L=Lpassed;
Th=Thpassed;  % this is T(1)
Rcon=Rconpassed;
Rload=Rloadpassed;

Tc=Tfluid;  % this is nominal T(N)

make_paramsAPL(L,factors);
%make_paramsSiGe(L,factors);

% xload is an equally spaced vector of positions
% Tload is a (not necessarily equally spaced) vector of temperatures
% params is a matrix giving the values of the physical parameters at
each position (row) and temperature (column)
load params x Tsamples params;
% x = depth of structure (cm)
% Tsamples = temperatures corresponding to parameter matrices (K)
KT=params(:,:,1);  % thermal conductivity with position and
temperature (W/cm/K)
alphaT=params(:,:,2);  % Seebeck coeff with position and temperature
(W/cm/K)
sigmaT=params(:,:,3);  % electrical conductivity with position
(1/Ohm/cm)
N=size(params,1);  % number of positions
M=size(params,2);  % number of temperatures

% make polynomial fits to material properties at each spatial position
KTfits=makeTfits(KT,Tsamples,3);
alphaTfits=makeTfits(alphaT,Tsamples,3);
sigmaTfits=makeTfits(sigmaT,Tsamples,3);

% take initial guess at temperature solution based on linear gradient
Tguess=linspace(Th,Tc,N);

% update values of physical parameters at each position based on
temperature information
% (there is an "update_param" and an "update_param_fast")
```

```
[K, K_hs, DK, alpha, alpha_hs, Dalpha, sigma,
sigma_hs]=update_param(Tguess, KTfits, alphaTfits, sigmaTfits);

% calculate results from bulk solver using averaged parameters
Kavg=1/mean(1./K);
alphaavg=sqrt(mean(alpha.^2))/2;
sigmaavg=1/mean(1./sigma);
[P,E,Tc]=fun_bulk(Th, Tfluid, H, L, Rcon, [Kavg alphaavg sigmaavg]);

% take new guess at temperature solution based on algebraic solution
Tguess=linspace(Th,Tc,N);

% update values of physical parameters at each position based on
temperature information
[K, K_hs, DK, alpha, alpha_hs, Dalpha, sigma,
sigma_hs]=update_param(Tguess, KTfits, alphaTfits, sigmaTfits);

A=1;  % area of TE device
Rcontacts=2*Rcon/A;  % ohms
Rseries=trapz(x,(1./sigma)/A)+Rcontacts;  % calculate resistance
if matched_load==1
    Rload=Rseries;  % assume matched load resistance
end

dx=x(2)-x(1);  % mesh spacing  (cm)
L=abs(x(end)-x(1));  % length of device (cm)

Kavg=1/mean(1./K);
dH=Th-Tc;
J_est=(Th-Tc)*mean(alpha)/(Rload+Rseries)/A;
GenP_est=(J_est*A)^2*Rload;
Pin_est=Kavg*dH*A/L+mean(alpha)*Th*J_est*A-(J_est*A)^2*Rseries/2;
Pout_est=Kavg*dH*A/L+mean(alpha)*Tc*J_est*A+(J_est*A)^2*Rseries/2;
Eff_est=GenP_est/Pin_est;


% try and make more accurate guess at internal power flux based on
temperature and parameters
%Qguess=mean(alpha)*J_est.*Tguess-K.*[(Tguess(2)-Tguess(1))
diff(Tguess)]/dx;
%for ind=1:length(Qguess)  % slightly better guess, accounting for
ohmic heat?
%   Qguess(ind)=Qguess(ind)+(ind-1)*J_est^2/sigma(ind)*dx;
%end
Qguess=linspace(Pin_est,Pout_est,N);

% initial guess at solution based on Ioffe-type bulk calculation
sol=[Tguess(2:end) Qguess J_est]';
T=Tguess;
Q=Qguess;
J=J_est;
norm=1;  % this will be the measure of convergence
```

```
numits=0;


while (norm>0.0001)&(numits<50)
   numits=numits+1;
   [K, K_hs, DK, alpha, alpha_hs, Dalpha, sigma,
sigma_hs]=update_param(T, KTfits, alphaTfits, sigmaTfits);  % find
parameters
   Rseries=trapz(x,(1./sigma)/A)+Rcontacts;  % calculate resistance
   if matched_load==1
        Rload=Rseries;  % assume matched load resistance
    end
   R=makeRhs(T, Q, J, dx, K_hs, alpha_hs, alpha, sigma_hs);  %
residual vector: size=(2*N-1,1)
    DR=makeDRhs(T, Q, J, dx, K_hs, alpha_hs, alpha, sigma_hs);  %
Jacobean matrix: size=(2*N-1,2*N-1)
   Du=-DR\R;  % compute the correction vector for Newton's method:
size=(2*N-1,1)
%   if (sol(end)+Du(end))<0
%      damping=0.5*sol(end)/Du(end);
%   else
%      damping=1;
%   end
   damping=1;  % parameter can be used to soften corrections, =1 is
standard newton method
   sol=sol+damping*Du;
   % update solution
   T=[T(1) sol(1:N-1)'];  % update whole-step temperatures
   Q=sol(N:(2*N-1))';  % update whole-step powers
   J=sol(end);

   T_mem(:,numits)=T';
   Q_mem(:,numits)=Q';
   R_mem(:,numits)=R;
   DR_mem(:,:,numits)=DR;
   Du_mem(:,numits)=Du;

% plot the solution
    do_calc_plots=0;
    if do_calc_plots

   figure(3);
   subplot(2,1,1);
   h=plot(x, T);
    xlabel('depth (cm)');
    ylabel('Temperature (K)');
    title(['Temperature vs. Depth: T_H=' num2str(Th) 'K, T_C='
num2str(Tc) 'K, J=' num2str(J) 'A/cm^2']);
    set(gca,'FontSize',14);
    set(gca,'LineWidth',2);
    set(h,'LineWidth',2);
    pause(0.01);
```

```
   % plot the solution
   subplot(2,1,2);
   h=plot(x, Q);
    xlabel('depth (cm)');
    ylabel('Power (W/cm^2)');
    title(['Power vs. Depth: T_H=' num2str(Th) 'K, T_C=' num2str(Tc)
'K, J=' num2str(J) 'A/cm^2']);
    set(gca,'FontSize',14);
    set(gca,'LineWidth',2);
    set(h,'LineWidth',2);
    pause(0.01);
    end


    norm=sum(R.^2);
    %norm=0;
    fprintf(1,'Iteration #%d: \t residual norm = %3.3e
\n',numits,norm);
end

R_int=Rseries;

T_hs=(T(1:end-1)+T(2:end))/2;
gradT_hs=diff(T)/dx;
x_hs=(x(1:end-1)+x(2:end))/2;
%sigma_hs=(sigma(1:end-1)+sigma(2:end))/2;
%alpha_hs=(alpha(1:end-1)+alpha(2:end))/2;
%K_hs=(K(1:end-1)+K(2:end))/2;
E_hs=J./sigma_hs+gradT_hs.*alpha_hs;
Pin=Q(1)*A;
Pout=Q(end)*A;
GenP=(J*A)^2*Rload;
Eff=GenP/Pin;

% try and estimate things with knowledge of the right Tc
Tc=T(end);
dH3=Th-Tc;
Kavg=1/mean(1./K);
Rseries=trapz(x,(1./sigma)/A)+Rcontacts;  % calculate resistance
J_est3=dH3*mean(alpha)/(Rload+Rseries)/A;
GenP_est3=(J_est3*A)^2*Rload;
Pin_est3=Kavg*dH3*A/L+mean(alpha)*Th*J_est3*A-(J_est3*A)^2*Rseries/2;
Pout_est3=Kavg*dH3*A/L+mean(alpha)*T(end)*J_est3*A+(J_est3*A)^2*Rserie
s/2;
Eff_est3=GenP_est3/Pin_est3;

eta_C=(Th-Tc)/Th;
Tave=(Tc+Th)/2;
eta=Eff;
% Z_est=(((eta_C+(eta*Tc/Th))/(eta_C-eta))^2-1)/Tave;  % infinite
staging result
```

```
Z_est=4/((Th-Tc)/eta-2*Th+(Th-Tc)/2);



fprintf(1,'\n');
fprintf(1,'\n');
fprintf(1,'TE unicouple numerical results:\n');
fprintf(1,'
_____________________________________________________________________
___________________ \n');
fprintf(1,'|                      | Predicted (unknown T(x)) |
Predicted (known T(x)) |  Simulated   |\n');
fprintf(1,'|-------------------------------------------------------------
--------------------------------|\n');
fprintf(1,'| Th-Tc    (deg K):    |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',dH,dH3,T(1)-T(end));
fprintf(1,'| matched load (Ohm):  |\t  \t  \t  |\t  \t \t  | %7.3e \t
|\n',Rseries);
fprintf(1,'| actual load (Ohm):   |\t  \t  \t  |\t  \t \t  | %7.3e \t
|\n',Rload);
fprintf(1,'| Current (A/cm^2):    |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',J_est,J_est3,J);
fprintf(1,'| Heat in (W):         |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',Pin_est,Pin_est3,Pin);
fprintf(1,'| Heat out (W):        |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',Pout_est,Pout_est3,Pout);
fprintf(1,'| Power generated (W): |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',GenP_est,GenP_est3,GenP);
fprintf(1,'| Efficiency (%%):      |\t %7.3f \t  |\t %7.3f \t  |\t
%7.3f  |\n',Eff_est*100,Eff_est3*100,Eff*100);
fprintf(1,'|_____________________________________________________________
______________________________|\n');
efficiency=Eff;
Powerout=GenP;
Powerout_E=Pout_est;
efficiency_E=Eff_est;


warning off;

c='r';

doplots=0;
if doplots

dobasic=1;
if dobasic

figure(1)
subplot(3,1,1);
hold on;
```

```
h=plot(x, K,'r');
xlabel('depth (cm)');
ylabel('Thermal Conductivity (W/cm/K)');
title('Thermal Conductivity vs. Depth');
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

subplot(3,1,2);
hold on;
h=plot(x, sigma,'r');
xlabel('depth (cm)');
ylabel('Electrical Conductivity (1/\Omega/cm)');
title('Electrical Conductivity vs. Depth');
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

subplot(3,1,3);
hold on;
h=plot(x, alpha,'r');
xlabel('depth (cm)');
ylabel('Seebeck Coefficient (V/K)');
title('Seebeck Coefficient vs. Depth');
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

end

figure(2);
subplot(2,1,1);
hold on;
h=plot(x_hs, gradT_hs,'r');
xlabel('depth (cm)');
ylabel('Gradient Temperature (K/cm)');
title(strcat('Gradient of Temperature vs. Depth: T_H=',num2str(Th),'K,
T_C=',num2str(Tc),'K'))
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

subplot(2,1,2);
hold on;
h=plot(x_hs, E_hs,'r');
xlabel('depth (cm)');
ylabel('Electric field (V/cm)');
```

```
title(strcat('Electric Field vs. Depth: T_H=',num2str(Th),'K,
T_C=',num2str(Tc),'K'))
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

figure(3);
subplot(2,1,1);
hold on;
h=plot(x, Q,'r');
xlabel('depth (cm)');
ylabel('Power Density (W/cm^2)');
title(strcat('Total Power Flux vs. Depth: T_H=',num2str(Th),'K,
T_C=',num2str(Tc),'K'))
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;

subplot(2,1,2);
hold on;
h=plot(x, T,'r');
xlabel('depth (cm)');
ylabel('Temperature (K)');
title(['Temperature vs. Depth: T_H=' num2str(Th) 'K, T_C=' num2str(Tc)
'K, J=' num2str(J) 'A/cm^2']);
set(gca,'FontSize',14);
set(gca,'LineWidth',2);
set(h,'LineWidth',2);
hold off;
warning on;

end

save record R_mem DR_mem Du_mem
```

```
function [P,E,Tc]=fun_bulk(Th, Tfluid, H, L, Rcon, params)
% ordering of params [K alpha sigma]

A=1;
if nargin<6
   load params
end

K=params(1);
alpha=params(2);
sigma=params(3);
R=L/sigma/A+2*Rcon/A;
sigma=L/R/A;
```

```
Z=alpha^2*sigma/K;
a=3*Z/8;
b=H*L/K+1-Z*Th/4;
c=-(Z/8*Th^2+Tfluid*H*L/K+Th);
Tc=(-b+sqrt(b^2-4*a*c))/(2*a);
V=alpha*(Th-Tc);
J=V/(2*R)/A;
P=(J*A)^2*R/A;
Pin=K*(Th-Tc)/L+alpha*Th*J-(J*A)^2*R/2;
E=P/Pin;
```

```
function [K, K_hs, DK, alpha, alpha_hs, Dalpha, sigma,
sigma_hs]=update_param(T, KTfits, alphaTfits, sigmaTfits);

N=length(T);  % number of positions the temperature is known at

% make some vectors to use for derivatives
DKvec=linspace((length(KTfits(1,:))-1),1,(length(KTfits(1,:))-1));
Dalphavec=linspace((length(alphaTfits(1,:))-
1),1,(length(alphaTfits(1,:))-1));

% find whole-step parameters
for ind=1:N
    K(ind)=polyval(KTfits(ind,:),T(ind));
    DKTfits=KTfits(ind,1:end-1).*DKvec; % derivative fit
   DK(ind)=polyval(DKTfits,T(ind));

   alpha(ind)=polyval(alphaTfits(ind,:),T(ind));
    DalphaTfits=alphaTfits(ind,1:end-1).*Dalphavec; % derivative fit
   Dalpha(ind)=polyval(DalphaTfits,T(ind));

    sigma(ind)=polyval(sigmaTfits(ind,:),T(ind));
end

% find half-step parameters
T_hs=(T(1:end-1)+T(2:end))/2;
for ind=1:N-1
    K_hs(ind)=polyval(KTfits(ind,:),T_hs(ind));
    DKTfits=KTfits(ind,1:end-1).*DKvec; % derivative fit
   DK_hs(ind)=polyval(DKTfits,T_hs(ind));

   alpha_hs(ind)=polyval(alphaTfits(ind,:),T_hs(ind));
    DalphaTfits=alphaTfits(ind,1:end-1).*Dalphavec; % derivative fit
   Dalpha_hs(ind)=polyval(DalphaTfits,T_hs(ind));

    sigma_hs(ind)=polyval(sigmaTfits(ind,:),T_hs(ind));
end
```

```
function Tfits=makeTfits(values, T, degree)
% index 1 of values refers to position, index 2 refers to temperature
% degree is what order polynomial will be computed

numtemp=size(values,2);
numpos=size(values,1);

Tfits=zeros(numpos,degree+1);
values_prev=values(1,:)*0;

for ind=1:numpos
   if (sum(values(ind,:)~=values_prev)==0)&(ind>1) % don't recompute
fit if you don't need to
      Tfits(ind,:)=Tfits(ind-1,:);
   else
      P=polyfit(T,values(ind,:),degree);
      Tfits(ind,:)=P;
   end
   values_prev=values(ind,:);
end
```

```
function DR=makeDRhs(T, Q, J, dx, K_hs, alpha_hs, alpha, sigma_hs);
% this makes the Jacobean matrix needed for Newton's method
% could probably make this much faster with spdiags command and
judicious linear algebra...

global A Rload Rcontacts Tfluid H

N=length(T);
DR=zeros(2*N,2*N);

row=1;
col=row;
DR(row,col)=alpha_hs(col)*J/dx;
DR(row,col+(N-1))=1/dx;
DR(row,col+(N-1)+1)=-1/dx;
DR(row,end)=2*J/sigma_hs(col)+alpha_hs(col)*(T(col+1)-T(col))/dx;

for row=2:N-1
    col=row;
    DR(row,col-1)=-alpha_hs(col)*J/dx;
    DR(row,col)=alpha_hs(col)*J/dx;
   DR(row,col+(N-1))=1/dx;
    DR(row,col+(N-1)+1)=-1/dx;
    DR(row,end)=2*J/sigma_hs(col)+alpha_hs(col)*(T(col+1)-T(col))/dx;
end

%row=N;
```

```
%col=row;
%DR(row,col-1)=-alpha_hs(col)*J/dx;
%DR(row,col+(N-2))=1/dx;
%DR(row,col+(N-2)+1)=-1/dx;
%DR(row,end)=2*J/sigma_hs(col)+alpha_hs(col)*(T(col+1)-T(col))/dx;


row=N;
col=row-N+1;
DR(row,col)=alpha_hs(col)*J/2-K_hs(col)/dx;
DR(row,col+(N-1))=-1/2;
DR(row,col+(N-1)+1)=-1/2;
DR(row,end)=alpha_hs(col)*(T(col+1)+T(col))/2;

for row=(N+1):(2*N-2)
    col=row-N+1;
    DR(row,col-1)=alpha_hs(col)*J/2+K_hs(col)/dx;
    DR(row,col)=alpha_hs(col)*J/2-K_hs(col)/dx;
   DR(row,col+(N-1))=-1/2;
    DR(row,col+(N-1)+1)=-1/2;
    DR(row,end)=alpha_hs(col)*(T(col+1)+T(col))/2;
end


%row=2*N-2;
%col=row-N+1;
%DR(row,col-1)=alpha_hs(col)*J/2+K_hs(col)/dx;
%DR(row,col+(N-1))=-1/2;
%DR(row,col+(N-1)+1)=-1/2;
%DR(row,end)=alpha_hs(col)*(T(col+1)+T(col))/2;


row=2*N-1;
DR(row,1:N-2)=J*A*(alpha_hs(2:end)-alpha_hs(1:end-1));
DR(row,N-1)=-J*A*alpha_hs(N-1);
DR(row,end)=-2*J*(Rload+Rcontacts)*(A^2)-
dx*A*sum(2*J./sigma_hs+alpha_hs.*diff(T)/dx);


row=2*N;
DR(row,N-1)=-H;
%DR(row,N-1)=-(J*alpha(N)+H);
DR(row,2*N-1)=1;
%DR(row,2*N)=-alpha(N)*T(N);
```

```
function R=makeRhs(T, Q, J, dx, K_hs, alpha_hs, alpha, sigma_hs);
% this makes the residual vector needed for Newton's method
```

```
global A Rload Rcontacts Tfluid H

N=length(T);
R=zeros(2*N,1);


R(1:N-1)=J^2./sigma_hs+alpha_hs*J.*(T(2:end)-T(1:end-1))/dx-(Q(2:end)-
Q(1:end-1))/dx;
%for row=1:N-1 %  temperature residual
%   R(row)=J^2./sigma_hs(row)+alpha_hs(row)*J*(T(row+1)-T(row))/dx-
(Q(row+1)-Q(row))/dx;
%end

R(N:(2*N-2))=alpha_hs*J.*(T(2:end)+T(1:end-1))/2-K_hs.*(T(2:end)-
T(1:end-1))/dx-(Q(2:end)+Q(1:end-1))/2;
%for row=1:N-1 %  power residual
%   R(row+(N-1))=alpha_hs(row)*J*(T(row+1)+T(row))/2-
K_hs(row)*(T(row+1)-T(row))/dx-(Q(row+1)+Q(row))/2;
%end

E=J./sigma_hs+alpha_hs.*diff(T)/dx;
R(2*N-1)=-(J*A)^2*(Rload+Rcontacts)-dx*J*A*sum(E);

R(2*N)=Q(N)-H*(T(N)-Tfluid);
%R(2*N)=(Q(N)-alpha(N)*J*T(N))-H*(T(N)-Tfluid);
```

```
function []=make_paramsAPL(L, factor_of_improvements)

if nargin==1
   factor_of_improvements=[1 1 1];
end

fK=factor_of_improvements(1);
falpha=factor_of_improvements(2);
fsigma=factor_of_improvements(3);


x=linspace(0,L,100);  % distance (cm)
Tsamples=linspace(216,1016,50);  % temperature (K)
params=zeros(length(x),length(Tsamples),3);
% 1 - K
% 2 - alpha
% 3 - sigma

for ind=1:floor(length(x))
   for Tind=1:length(Tsamples)
      params(ind,Tind,1)=0.0393*fK; % (W/cm/K)
      params(ind,Tind,2)=247e-6*falpha;%*(1+5*(ind-1)/(length(x)));%
(V/K)
```

```
      params(ind,Tind,3)=485*fsigma; % (1/Ohm/cm)
   end
end

save params x Tsamples params;
```

****

Chapter 3 code:

1) Heat spreading and thermal quadrupole code

2) Microchannel channel simulations

3) TE generator + Microchannel system level simulations

Heat spreading and thermal quadrupole code

```
function [Zth,
Mspread]=getRspread_dynamic(Adev,Asub,Lsub,ksub,alphasub,h,f,J1sols)
% function [Zth, Mspread]=getRspread_dynamic
(Adev,Asub,Lsub,ksub,alphasub,h,f,J1sols)
%
% outputs:
% Zth - total complex thermal impedance of substrate including
spreading,
% substrate thickness, and heat sink.  [deg C/Watt]
% Mspread - complex thermal quadrupole matrix for the spreading and
%   substrate thermal transport (does not include heat sink impedance)
%
% inputs:
% Adev - area of small circular heat emitter [cm^2]
% Asub - area of larger circular heat spreader [cm^2]
% Lsub - thickness of larger heat spreader [cm]
% ksub - thermal conductivity of heat spreader [W/cm/K]
% alphasub - thermal diffusivity of heat spreader [cm^2/s]
% h - heat transfer coefficient at bottom of spreader [W/cm^2/K]
% f - frequency of interest (2*pi*f=imag(s) in Laplace domain)
% J1sols - vector containing the zeros of the Bessel function of order
1
%   (see function besselzero.m)
%
% note: For the solution used here, the heat flux into the substrate
and
% the heat transfer out of the substrate are assumed uniform.
%
```

```
if nargin<8
    J1sols=besselzero(1,300,1);
end

r0=sqrt(Adev/pi);
R=sqrt(Asub/pi);
d=Lsub;
k=ksub;

s=j*2*pi*f;
S=pi*R^2;
lam=sqrt(s/alphasub);
A=cosh(lam*d);
B=sinh(lam*d)/(k*lam*S);
C=k*lam*S*sinh(lam*d);
Bi=h*R/k;

alphsols=J1sols/R;  % first few bessel zeros (scaled)
gamsols=sqrt(alphsols.^2+s/alphasub);
Fterm=(4*besselj(1,alphsols*r0).^2)./(k*S*r0^2*alphsols.^2.*gamsols.*b
esselj(0,alphsols*R).^2).*((tanh(gamsols*d)+(alphsols*R./Bi))./(1+(alp
hsols*R./Bi).*tanh(gamsols*d)));
sumFterm=sum(Fterm);

Z11=A+C*sumFterm;
Z12=B+A*sumFterm;
Z21=C;
Z22=A;

M11=Z22;
M12=-Z12;
M21=-Z21;
M22=Z11;
Mspread=[M11 M12 ; M21 M22];

hs=[pi*R^2*h -1];
Vout2=hs*Mspread;
Zth=-Vout2(2)./Vout2(1);

% the terms below are the "T-network" thermal impedances for the
spreader
%Z1=sumFterm+1/(k*lam*S)/tanh(lam*d)-1/(k*lam*S)/sinh(lam*d);
%Z2=1/(k*lam*S)/tanh(lam*d)-1/(k*lam*S)/sinh(lam*d);
%Z3=1/(k*lam*S)/sinh(lam*d);
%Zth3=Z1+1/(1/Z3+1/(Z2+1/(pi*R^2*h)));
```

```
function x=besselzero(n,k,kind)
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%
%
% besselzero.m
%
% Find first k positive zeros of the Bessel function J(n,x) or Y(n,x)
% using Halley's method.  Set kind=1 for J(n,x), kind=2 for Y(n,x).
%
% Written by: Greg von Winckel - 01/25/05
% Contact: gregvw(at)chtm(dot)unm(dot)edu
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%

k3=3*k;

x=zeros(k3,1);

for j=1:k3

    % Initial guess of zeros
    x0=1+sqrt(2)+(j-1)*pi+n+n^0.4;

    % Do Halley's method
    x(j)=findzero(n,x0,kind);

    if x(j)==inf
        error('Bad guess.');
    end

end

x=sort(x);
dx=[1;abs(diff(x))];
x=x(dx>1e-8);

x=x(1:k);

function x=findzero(n,x0,kind)

n1=n+1;     n2=n*n;

% Tolerance
tol=1e-12;

% Maximum number of times to iterate
MAXIT=100;

% Initial error
err=1;
```

```
iter=0;

while abs(err)>tol & iter<MAXIT

    switch kind
        case 1
            a=besselj(n,x0);
            b=besselj(n1,x0);
        case 2
            a=bessely(n,x0);
            b=bessely(n1,x0);
    end

    x02=x0*x0;

    err=2*a*x0*(n*a-b*x0)/(2*b*b*x02-a*b*x0*(4*n+1)+(n*n1+x02)*a*a);

    x=x0-err;
    x0=x;
    iter=iter+1;

end

if iter>MAXIT-1
    warning('Failed to converge to within tolerance. ',...
            'Try a different initial guess');
    x=inf;
end
```

```
function [Rspread,Rthbulk]=getRspread(Adev,Asub,Lsub,ksub,h,calc_type)
% function
[Rspread,Rthbulk]=getRspread(Adev,Asub,Lsub,ksub,h,calc_type)
% outputs:
% Rspread - total resistance of substrate including spreading and
contact effects [deg C/Watt]
% Rthbulk - appears to be the substrate thermal resistance
% inputs:
% Adev - area of small circular heat emitter
% Asub - area of larger circular heat spreader
% Lsub - thickness of larger heat spreader
% k_sub - thermal conductivity of heat spreader
% h - heat transfer coefficient at bottom of spreader
% calc_type=1 is the Song, Lee, Van Au result for spreading resistance
% calc_type=2 is the Yovanovich Antonetti result for spreading
resistance

if nargin<6
   calc_type=1;  % assume Song et al calculation if none given
end

if nargin<5
```

```
   h=1e8;  % assume perfect heat sink if none is given
end

rdev=sqrt(Adev/pi);
rsub=sqrt(Asub/pi);
tau=Lsub/rsub;

Rr=rdev./rsub;
Rspread2=(1-
1.410*Rr+0.344*Rr.^3+0.043*Rr.^5+0.034*Rr.^7)./(4*ksub.*rdev)+tau.*Rr.
/ksub./rdev/pi;

R_0=1./h/pi/(rsub.^2);
lambdac=pi+1./(sqrt(pi)*Rr);
Bi=1/pi./ksub./rsub/R_0;
phi=(tanh(lambdac.*tau)+lambdac./Bi)./(1+lambdac./Bi.*tanh(lambdac.*ta
u));
psi=Rr.*tau/sqrt(pi)+(1/2)*phi.*(1-Rr).^(3/2);
Rspread1=psi/sqrt(pi)./ksub./rdev;
Rthbulk=Rr.*tau/sqrt(pi)/sqrt(pi)./ksub./rdev;

switch calc_type
case{1}
   Rspread=Rspread1;
case{2}
   Rspread=Rspread2;
end
```

Microchannel channel simulations

```
% This is a finite difference solver for fully developed laminar flow
in a rectangular
% channel.  The flow is solved analytically, and the temperature
through
% an iterative finite difference scheme.

clear all;


b=100e-6;  % channel height z (m)
a=100e-6;  % channel width y (m)
% s=40e-6; % channel spacing (fin width) (m)

Numy=100; Numz=200; % number of segments
dx=2e-6;
A=a*b;
Dh=4*A/(2*a+2*b); % def'n of hydraulic diameter

rho1=998; rho2=6363.2; % densities for 1-water 2-Ga68 In20 Sn12
```

```
k1=0.606; k2=39; % thermal conductivities
cp1=4181; cp2=365.813; % heat capacities
nu1=0.960*1e-6; nu2=0.34809*1e-6; % kinematic viscosities
mu1=nu1*rho1; mu2=nu2*rho2; % dynamic viscosities
Pr1=cp1*mu1/k1; Pr2=cp2*mu2/k2; % Prandtl numbers

rho=rho2;
mu=mu2;
cp=cp2;
k=k2;
qs=10/1e-4;
Tb=320;

dPdx=1*1e5/1e-2; % first num is bar/cm

ord=20;
[B,k_]=calc_Four_coeff2(dPdx,mu,a,b,ord);
[u,y,z]=make_u_Four2(a,b,Numy,Numz,dPdx,mu,B,k_); % these have Num+1
points

dy=y(2)-y(1);
dz=z(2)-z(1);

figure(1);
%uview=[[u zeros(size(u,1),1)] ; zeros(1,size(u,2)+1)]; % fix for
matlab surf peculiarity
surf(z,y,u);
view(3);
xlabel('z');
ylabel('y');
zlabel('u');

ub=trapz(z,trapz(y,u))/A
exact_ub=dPdx*2*Dh^2/mu/57 % exact for square tube
G=rho*ub; % mass flow velocity
Rey=G*Dh/mu % Reynolds number
f=dPdx*Dh*2/rho/ub^2 % Darcy friction factor
Rey_f=Rey*f

pause;
% now calculate the temperature distribution

T=ones(size(u,1)+2,size(u,2)+2)*320;

figure(2);

epsilon=inf;

warning off;
while epsilon>0.000001

Told=T;
```

```
Tintold=T(2:end-1,2:end-1);

% first set the fake boundaries to the right temperatures

% set fake wall temps
T(1,3:end-2)=qs/k*dy+Tintold(1,2:end-1);
T(end,3:end-2)=qs/k*dy+Tintold(end,2:end-1);
T(3:end-2,1)=qs/k*dz+Tintold(2:end-1,1);
T(3:end-2,end)=qs/k*dz+Tintold(2:end-1,end);

% set fake corner temps
T(1,2)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(1,1);
T(2,1)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(1,1);
T(1,end-1)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(1,end);
T(2,end)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(1,end);
T(end-1,1)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(end,1);
T(end,2)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(end,1);
T(end,end-1)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(end,end);
T(end-1,end)=qs/k*(dz+dy)/(dz/dy+dy/dz)+Tintold(end,end);


% set the source term (total amount of power flowing out of the liquid
%       exactly equal to the power flowing in
dTdx=(2*(Numy)*qs*dy+2*(Numz)*qs*dz)/(sum(sum(u))*rho*cp*dz*dy);

Tzplus=T(2:end-1,1:end-2);
Tzminus=T(2:end-1,3:end);
Typlus=T(1:end-2,2:end-1);
Tyminus=T(3:end,2:end-1);

Tintnew=(-
u*rho*cp/k*dz*dy*dTdx+dy/dz*(Tzplus+Tzminus)+dz/dy*(Typlus+Tyminus))/(
2*(dy/dz+dz/dy));

w=1.00; % over-relaxation weight
%wmat=ones(size(dTint));
dTint=Tintnew-Tintold;
%dTint_norm=dTint/max(max(abs(dTint))).^2;
dTintguess=w*(dTint);

dTguess=[zeros(1,Numz+3) ; [zeros(Numy+1,1) dTintguess
zeros(Numy+1,1)] ; zeros(1,Numz+3)];
T=T+dTguess;
Tint=T(2:end-1,2:end-1);

Tbmeas=trapz(z,trapz(y,u.*Tint))/ub/A;
%Tbmeas=sum(sum(u.*Tint*dy*dz))/ub/A;
dTb=Tb-Tbmeas;
T=T+dTb;

epsilon=max(max(abs(1-T./Told)));
errstr=sprintf('error= %e',epsilon);
```

```
if epsilon==NaN
   epsilon=inf;
end

surf(z,y,Tint);
view(3);
xlabel('z');
ylabel('y');
zlabel('u');
title(errstr);

view(3);
pause(0.01);

end
warning on;


%Twall=(mean(Tint(1,:))+mean(Tint(end,:))+mean(Tint(:,1))+mean(Tint(:,
end)))/4
Twall=[Tint(1,:) Tint(end,:) Tint(:,1)' Tint(:,end)'];
%Twall=[(Tint(1,:)+T(1,2:end-1))/2 (Tint(end,:)+T(end,2:end-1))/2
((Tint(:,1)+T(2:end-1,1))/2)' ((Tint(:,end)+T(2:end-1,end))/2)'];
Twall_mean=mean(Twall)
h=mean(qs./(Twall-Tb))  % pays to do the averaging right (denom)
Nu=h*Dh/k

Nupois1=3.6
hpois1=Nupois1*k/Dh % local heat transfer coeff
```

```
function [u,y,z]=make_u_Four2(a,b,Numy,Numz,dPdx,mu,B,k)

dy=a/Numy;
dz=b/Numz;
y=0:dy:a;
z=0:dz:b;
u=zeros(Numy+1,Numz+1);
for k_=1:length(B)
    for y_=1:size(u,1)

u(y_,:)=u(y_,:)+B(k_)*sin(k(k_)*pi/a*y(y_))*cosh(k(k_)*pi/a*(z-b/2));
    end
end

for y_=1:size(u,1)
    u(y_,:)=u(y_,:)+(dPdx/2/mu)*(a*y(y_)-y(y_)^2);
end
```

```
function [B,k]=calc_Four_coeff2(dPdx,mu,a,b,ord)
```

```
% calculates the Fourier components for the vel dist in a sq chan
% meant to be used with MuCh_Four2.m

if nargin<5
    ord=50;
end

k=1:2:ord*2-1;
k_=1:length(k);
B(k_)=2*dPdx*a^2/mu/pi^3./k.^3./cosh(k*pi*b/2/a).*(cos(k*pi)-1);
```

TE generator + Microchannel system level simulations

```
% this code solves for the performance of a thermoelectric generator
% integrated with a microchannel cooling system.
%
% it sweeps the flow velocity (Reynolds number), but can be configured
to
% sweep through whatever.

clear all;

% boundary conditions
Th=273+227; % TE hot side temperature [K]
Tin=300; % channel inlet temperature [K]

% geometry
L=10e-3;  % length of mu-channels [m]
nc=36;  % number of parallel microchannels

b=1000e-6;  % height of mu-channel [m]
a=300e-6;  % width of mu-channel [m]
s=100e-6;   % fin thickness [m]
W=(a+s)*nc;  % width of mu-channel platform [m]
A=L*W;  % area of heat sink platform

t=100e-6;  % thickness of the substrate above channels, below TE [m]
k_sub=1.3*100;  % thermal conductivity of substrate

fA_TE=1;%0.6;  % fractional area of the heat sink covered by TE legs
nTE=18; % number of TE couples on die surface (1 couple = 2 elements)
L_TE=20e-6;  % length of the TE elements
A_TE=A*fA_TE;  % total area of TE devices on heat sink;
A_TE1=A_TE/2/nTE; % area of one TE element

% cooling fluid
[density,spec_heat,th_cond,visc]=get_thermophys(307,'water');
```

```
coolant_vec=[density,spec_heat,th_cond,visc];

% performance numbers
eta_pump=0.55; % fractional pump efficiency (1 -> perfect pump)
alpha=(190e-6+230e-6)/2; % effective Seebeck for n/p element [V/K]
sigma=(741+571)/2*100; % effective electrical conductivity for n/p
element [1/Ohm/m]
kappa=(0.014+0.012)/2*100; % effective thermal conductivity for n/p
element [W/m/K]

Rth_TE=L_TE/A_TE/kappa;  % thermal resistance of all of the TE
elements
R_TE=2*nTE*L_TE/A_TE1/sigma;  % series resistance of all of the TE
elements

%Th_vec=linspace(Tin+1,Th,20);
u=linspace(0.01,10,100);

for k=1:length(u)

[Pgen_net(k),Pgen(k),pump_power(k),dP(k),dP_dx(k),Eff(k),Qout(k),I(k),
RL(k),Tc(k),Tm(k),Tf(k),Re(k),Nu(k),fD(k),fin_eff(k),Rspread(k),h(k),h
eff(k)]=do_TE_mu_chan(Th,Tin,u(k),L,b,a,s,t,k_sub,nc,eta_pump,alpha,si
gma,kappa,fA_TE,nTE,L_TE,0,coolant_vec);
end


I=(2*nTE*alpha*(Th-Tc))/(2*R_TE);  % assuming impedance matching
Q_throughTE_sigma=(1/2)*(I.^2)*R_TE;
Q_throughTE_kappa=(Th-Tc)/Rth_TE;
Q_throughTE_alpha=Tc*alpha.*I*2*nTE;
Q_throughTE_tot=Q_throughTE_sigma+Q_throughTE_kappa+Q_throughTE_alpha;
Q_throughfins=(Tm-Tf).*heff*A;
Q_throughspread=nTE*2*(Tc-Tm)./Rspread;

% do plots of temperatures
figure(1);
plot(u,Th*ones(size(Re)),'r',u,Tc,'g',u,Tm,'b',u,Tf,'k')
title('Key system temperatures');
xlabel('Mean flow velocity (m/s)');
ylabel('Temperature (K)');
legend('Hot side of TE','Cold side of TE','Top of channels','Mean
fluid temperature');
prep_ppt_plot;

% do plots of practical pump issues
dPatm=dP/1e5;
figure(2);
plot(u,dPatm,'b',u,dP_dx*L/1e5,'r');
title('Entrance, exit, and 90 degree bend effects');
xlabel('Mean flow velocity (m/s)');
ylabel('Total pressure drop (atm)');
```

```
legend('Pressure drop','Ideal pressure drop');
prep_ppt_plot;

% do plots of channel fluid dynamics
clf;
figure(3);
[axyy h1yy h2yy]=plotyy(Re,fD,Re,Nu,'semilogy');
hold on;
xlabel('Re (Reynolds number = \rhouD_{h}/\mu)');
hylabel1=get(axyy(1),'Ylabel');
hylabel2=get(axyy(2),'Ylabel');
set(hylabel1,'String','f_{D} = Darcy friction factor');
set(hylabel2,'String','Nu = Nusselt number');
hold off;
%prep_ppt_plot;

% do plots of practical pump issues
figure(4);
plot(u,ones(size(u)),'k:',u,h/1e4,'b',u,((1./heff+t/k_sub).^(-
1))/1e4,'r');
title('Origins of enhanced heat sinking');
xlabel('Mean flow velocity (m/s)');
ylabel('h (W/cm^2/K)');
legend('macroscale baseline','channel wall h','Effective h (after
substrate+fins)');
prep_ppt_plot;

% do plots of system performance
figure(5);
[axyy h1yy h2yy]=plotyy(u,Pgen_net/A/1e4,u,Eff*100);
hold on;
axes(axyy(1));
plot(u,Pgen/A/1e4,'b-.');
title('System performance specs');
xlabel('Mean flow velocity (m/s)');
hylabel1=get(axyy(1),'Ylabel');
hylabel2=get(axyy(2),'Ylabel');
set(hylabel1,'String','Net generated power density (W/cm^2)');
set(hylabel2,'String','Efficiency (%)');
hold off;
axes(axyy(2));
%legend(axyy(1),'Net power','TE power');
%legend(axyy(2),'Efficiency','Location','SouthWest');

% do plots of system performance 2
figure(6);
plot(Eff*100,Pgen_net/A/1e4,'b');
title('Net power density and efficiency with Re as a parameter');
xlabel('Efficiency (%)');
ylabel('Power density (W/cm^2)');
prep_ppt_plot;
```

```
% this is just a check that power is conserved through the whole model
figure(7);
plot(Re,Q_throughTE_kappa,'k.',Re,Q_throughTE_sigma,'k:',Re,Q_throughT
E_alpha,'k-
.',Re,Q_throughfins,'go',Re,Qout,'bx',Re,Q_throughspread,'yh',Re,Q_thr
oughTE_tot,'rs');
xlabel('Re (Reynolds number = \rhouD_{h}/\mu)');
ylabel('Heat though device (W)');
title('Energy conservation check');
%prep_ppt_plot;

flow_rate=u*a*b*nc*1e6*60;  % mL/minute
figure(8);
[axyyb,h1yyb,h2yyb]=plotyy(flow_rate,dPatm,flow_rate,pump_power);
title('Pump needs');
xlabel('Flow rate (mL/min)');
hylabel1=get(axyyb(1),'Ylabel');
hylabel2=get(axyyb(2),'Ylabel');
set(hylabel1,'String','Total pressure drop (atm)');
set(hylabel2,'String','Total pump electrical power (W)');
%legend([hylabel1 hylabel2],'Pressure drop','Ideal pressure drop');

[max_power,max_power_ind]=max(Pgen_net);
max_power=max_power/1e4/A
max_efficiency=max(Eff)*100
eff_at_max_power=Eff(max_power_ind)*100
```

```
function
[density,spec_heat,th_cond,visc]=get_thermophys(T,coolant_type)
% function
[density,spec_heat,th_cond,visc]=get_thermophys(T,coolant_type)
%
% This function returns vectors with the thermophysical properties of
%   coolants evaluated at the temperatures of interest.
% From Incropera (1999)
%
% Outputs:
% density [kg/m^3]
% spec_heat - specific heat [J/kg/K]
% th_cond - thermal conductivity [W/m/K]
% visc - dynamic viscosity [N*s/m^2]
%
% Inputs:
% T - vector of temperatures of interest [K]
% coolant_type - either 'FC-72', 'FC-77', 'water', or 'metal'
%   (FC coolants are Fluorinert-type, metal is [Ga68 In20 Sn12])

if nargin<2
    coolant_type='water';
end
```

```
switch upper(coolant_type)
    case 'WATER'
        TC=T-273;
        density=(999.8396+18.22494*TC-7.92221e-3*TC.^2-5.54485e-
5*TC.^3+1.49756e-7*TC.^4-3.93295e-10*TC.^5)./(1+1.81597e-2*TC);
        spec_heat=8958.9-40.535*T+0.11243*T.^2-1.0138e-4*T.^3;
        th_cond=-0.58166+6.3555e-3*T-7.9643e-6*T.^2;
        visc=2.414e-5*10.^(247.8./(T-140));

    case 'FC-72'
        density=2453-2.61*T;
        spec_heat=585+1.550*T;
        th_cond=0.090-1.10e-4*T;
        visc=1.0017e-2-5.12375e-5*T+6.7252e-8*T.^2;
    case 'FC-77'
        density=2507-2.45*T;
        spec_heat=579+1.572*T;
        th_cond=0.0842-6.302e-5*T-2.600e-8*T.^2;
        visc=density*(1.4347e-5-7.7391e-8*T+1.0725e-10*T.^2);
    case 'METAL'
        density=6363.2*ones(size(T));
        spec_heat=139.068*ones(size(T));
        th_cond=39*ones(size(T));
        visc=2.2150e-3*ones(size(T));
end
```

```
function
[Pgen_net,Pgen,pump_power,dP,dP_dx,Eff,Qout,I,RL,Tc,Tm,Tf,Re,Nu,fD,fin
_eff,Rspread,h,heff]=do_TE_mu_chan(Th,Tin,u,L,b,a,s,t,k_sub,nc,eta_pum
p,alpha,sigma,kappa,fA_TE,nTE,L_TE,e_Dh,coolant_vec)
% function
[Pgen_net,Pgen,pump_power,dP,dP_dx,Eff,Qout,I,RL,Tc,Tm,Tf,Re,Nu,fD,fin
_eff,Rspread,h,heff]=do_TE_mu_chan(Th,Tin,u,L,b,a,s,t,k_sub,nc,eta_pum
p,alpha,sigma,kappa,fA_TE,nTE,L_TE,e_Dh,coolant_vec)
%
% This code calculates the performance of a TE module integrated with
a silicon
%   microchannel cold plate pumped with water coolant.
% It calls: mu_ch_fun.m, get_fin_eff.m, get_fD.m, get_Nu.m,
getRspread.m
%
% ***  All units should be in mks ***
%
% output parameters:
% Pgen_net - net power generated -- power dropped in load minus power
spent
%   on pump [W]
% Pgen - power generated in load resistor [W]
```

```
% pump_power - power spent pumping, including all entrance and exit
effects, and pump eff, for all channels [W]
% dP - pressure drop, including all entrance and exit effects [Pa]
% dP_dx - hydraulic, fully developed pressure drop per unit length of
%   channel, neglecting entrance and exit effects [Pa/m]
% Eff - overall thermal efficiency of generator
% Qout - total heat pumped into channels [W]
% I - closed circuit current [A]
% RL - load resistance (assumed to be matched to total elt res) [Ohm]
% Tc - temperature at cold side of generator [K]
% Tm - temperature at top of channels, after heat spreading [K]
% Tf - mean temperature of fluid in channel [K]
% Re - Reynolds number of channel flow
% Nu - Nusselt number of channel flow
% fD - Darcy (a.k.a. Moody) friction factor
% fin_eff - fractional efficiency of fin (from Kandlikar/Upadhye
paper)
% Rspread - spreading thermal resistance seen by a TE element into the
fins
% h - heat transfer coefficient at walls of mu-channel [W/m^2/K]
% heff - effective heat transfer coefficient at top of fins [W/m^2/K]
%
%
% input parameters:
% Th - hot side temperature [K]
% Tin - inlet fluid temperature [K]
% u - flow velocity [m/s]
% L, b, a - length, height, width of mu-channel [m]
% s - fin thickness
% t - substrate thickness (distance from bottom of TE to top of
channel [m]
% k_sub - substrate thermal conductuctance [W/m/K]
% nc - number of channels
% eta_pump - fractional pump efficiency (1 -> perfect pump)
% alpha - mean Seebeck coeff of n and p legs [V/K]
% sigma - mean elec conductivity of n and p legs [1/Ohm/cm]
% kappa - mean thermal conductivity of n and p legs [W/cm/K]
% fA_TE - fraction of the heat sink covered by TE (assuming equal
spaced
%   elts)
% nTE - number of TE couples (1 couple = 2 elements)
% L_TE - length of TE legs [m]
% e_Dh = surface roughness/hydraulic diameter (not used now)
% coolant_vec = physical parameters of coolant, a [1,4] vector of:
%   <mass density [kg/m^3]>
%   <mass heat capacity [J/kg/K]>
%   <conductivity [W/m/K]>
%   <dynamic viscosity [N s/m^2]>

if nargin<19
% assume water as coolant (no temperature fits)
    coolant_vec=[999 4188 0.588 0.00114];
```

```
end

if nargin<18
    e_Dh=0;
end

rho_f=coolant_vec(1); % mass density [kg/m^3]
cp_f=coolant_vec(2); % mass heat capacity [J/kg/K]
k_f=coolant_vec(3); % conductivity [W/m/K]
mu_f=coolant_vec(4); % dynamic viscosity [N s/m^2]

% geometry/substrate
Dh=4*a*b/(2*(a+b));  % hydraulic diameter [m]
width_tot=nc*(s+a);  % width of the cooling wafer [m]
A=width_tot*L;  % total area of cooling wafer [m^2]

% channel flow characteristics
mdot=rho_f*u*a*b;  % mass flow through one channel [kg/s]
[h, total_psi, pump_power, dP, dP_dx, Re, Pr, fD,
Nu]=calc_mu_ch(u,L,b,a,coolant_vec,e_Dh);

% fin characteristics
fin_eff=get_fin_eff(a, b, s, L, k_sub,h);
Aw=(2*fin_eff*b+a)*L*nc;  % effective area of heat transfer to water
heff=h*Aw/A; % effective heat transfer coefficient from heatsink to
water

% system calculations
pump_power=pump_power/eta_pump*nc;  % power used to drive pump
total_pressure=dP_dx*L; % in pascals

% TE geometry calculations

A_TE=fA_TE*A;  % area covered by all the TE n+p couples
A_TE1=A_TE/(2*nTE); % area of 1 TE element (n or p, assume they are
same for now...)
A1=A/(2*nTE); % area on heat sink below 1 TE element
[Rspread,Rthbulk]=getRspread(A_TE1,A1,t,k_sub,heff,1); % calculate
thermal resistance from TE to sink (spreading)

% TE performance calculations

% Th (top of TE), Tc (bottom of TE), Tm (right above water in finned
heat sink), Tf (avg water temp)

R_elt=L_TE/sigma/A_TE1;  % element electrical resistance
RL=2*nTE*R_elt;  % assume matched load
syms I Tc Tm Tf

eqncurrent=(2*nTE*alpha*(Th-Tc))/(2*nTE*R_elt+RL)-I;
eqntop=(Tc-Tm)/Rspread-((Th-
Tc)*kappa*A_TE1/L_TE+I*alpha*Tc+(1/2)*R_elt*I^2);
```

```
eqnbot=(Tc-Tm)/Rspread-(Tm-Tf)*heff*A1;
eqnfluid=(Tm-Tf)*heff*A/nc/mdot/cp_f-(2*Tf-2*Tin);
S4=solve(eqncurrent,eqnbot,eqntop,eqnfluid,I,Tm,Tc,Tf);
Tm_vec=eval(S4.Tm);
Tc_vec=eval(S4.Tc);
Tf_vec=eval(S4.Tf);

% select valid solution by checking that temperatures are well-ordered
found_sol=0;
if length(Tm_vec>=1)
    for ind=1:length(Tm_vec)
        if
(Th>Tc_vec(ind))&(Tc_vec(ind)>Tm_vec(ind))&(Tm_vec(ind)>Tf_vec(ind))&(
Tf_vec(ind)>Tin)
            if found_sol==1
                disp('Error: multiple temperature solutions found');
            end
            Tm=Tm_vec(ind);
            Tc=Tc_vec(ind);
            Tf=Tf_vec(ind);
            I=eval(S4.I(ind));
            found_sol=1;
        end
    end
end
if found_sol==0
    disp('Error: no well ordered temperatures found');
end

Tout=Tf*2-Tin;
V=I*RL;
Pgen=I*V;
Pgen_net=Pgen-pump_power;
Qout=(Tout-Tin)*cp_f*mdot*nc;
Eff=Pgen_net/(Qout+Pgen);
```

```
function [h, total_psi, pump_power, dP, dP_dx, Re, Pr, fD,
Nu]=calc_mu_ch(u,L,b,a,coolant_vec,e_Dh)
%
% function [h, total_psi, pump_power, dP, dP_dx, Re, Pr, fD,
Nu]=calc_mu_ch(u,L,b,a,coolant_vec,e_Dh)
% Calculates flow and heat transfer in microchannel using empirical
%   correlations.
%
% outputs:
% h = heat transfer coefficient at wall of channel [W/m^2/K]
% total_psi = total pressure drop across channel [psi] (note 14.2
psi=1 atm)
% pump_power = total hydraulic power needed to pump channel [W]
```

```
% dP = pressure drop across channel including all entrance and exit
effects [Pa]
% dP_dx = fully developed pressure drop per unit length [Pa/m]
% Re = Reynolds number
% Pr = Prandtl number
% fD = Darcy (aka Moody) friction factor
% Nu = Nusselt number
%
% inputs:
% u = fluid velocity [m/s]
% L, b, a - length, height, width of mu-channel [m]
% coolant_vec = physical parameters of coolant, a [1,4] vector of:
%   <mass density [kg/m^3]>
%   <mass heat capacity [J/kg]>
%   <conductivity [W/m/K]>
%   <dynamic viscosity [N s/m^2]>
% e_Dh = surface roughness/hydraulic diameter

if nargin<6
    e_Dh=0;
end

if nargin<5
% assume water as coolant (no temperature fits)
    coolant_vec=[999 4188 0.588 0.00114];
end


rho_f=coolant_vec(1); % mass density [kg/m^3]
cp_f=coolant_vec(2); % mass heat capacity [J/kg]
k_f=coolant_vec(3); % conductivity [W/m/K]
mu_f=coolant_vec(4); % dynamic viscosity [N s/m^2]

% geometry/substrate
Dh=4*b*a/(2*(b+a));  % hydraulic diameter [m]

% channel flow characteristics
Re=rho_f*u*Dh/mu_f; % Reynolds number of the flow (determines
laminar/turbulent)
zplus=0.08*Dh*Re; % entrance length estimate (p72 Incropera) [m]
Pr=cp_f*mu_f/k_f; % Prandtl number (diffusive momentum
tranfer/diffusive heat transfer)
fD=get_fD(Re, b, a, L); % Darcy friction factor
dP_dx=1/2*rho_f*u^2/Dh*fD; % Pressure drop per unit length of channel
neglecting entrance/exit effects

% bend and entrance and exit losses (developing flow is a separate
issue!)
K90=1.2; % loss coefficient for 90 degree bend
Kc=0.75; % loss coefficient for entrance of channel (estimate)
Ke=0.75; % loss coefficient for exit of channel (estimate)
```

```
%K90=0.0; % loss coefficient for 90 degree bend
%Kc=0.0; % loss coefficient for entrance of channel (estimate)
%Ke=0.0; % loss coefficient for exit of channel (estimate)

dP=1/2*rho_f*u^2*(2*K90+Kc+Ke+fD*L/Dh);
pump_power=b*a*u*dP;  % Hydraulic pump power (w/o effects of pump
efficiency)

% thermal flow characteristics
Nu=get_Nu(Re, Pr, b, a, L);
h=Nu*k_f/Dh;

% system calculations
total_psi=0.000145*dP;
```

```
function [fD, fully_developed_frac]=get_fD(Re, b, a, L);
% function [fD, fully_developed_frac]=get_fD(Re, b, a, L);
% Calculate Darcy/Moody friction factor for developing flow
%   (uses Phillips 1988 results, from Incropera)
% Averages the result over the channel length.
%
% outputs:
% fD=Darcy friction factor
% fully_developed_frac=fraction of the length of the channel over
which the
%   flow is considered fully developed
%
% inputs:
% Re=Reynolds number, calculated with hydraulic diameter (vector)
% b=channel height [m]
% a=channel width [m]
% L=channel length [m]

aspect=min([a b])/max([a b]);

% this one I don't use, but it's nice to have for comparison...
table_lam=[             % pg 77 of Incropera (fully developed laminar)
0.0 24.00;
0.1 21.17;
0.2 19.07;
0.3 17.51;
0.4 16.37;
0.5 15.55;
0.6 14.98;
0.7 14.61;
0.8 14.38;
0.9 14.26;
1.0 14.23];

% this one is used for the laminar correlations
```

```
% z A=1.0 A=0.5 A=0.2 A<0.1
table_dev_lam=[          % pg 78 of Incropera (developing flow)
0.001   111.0   111.0   111.0   112.0;
0.003   66.0    66.0    66.1    67.5;
0.005   51.8    51.8    52.5    53.0;
0.007   44.6    44.6    45.3    46.2;
0.009   39.9    40.0    40.6    42.1;
0.010   38.0    38.2    38.9    40.4;
0.015   32.1    32.5    33.3    35.6;
0.020   28.6    29.1    30.2    32.4;
0.030   24.6    25.3    26.7    29.7;
0.040   22.4    23.2    24.9    28.2;
0.050   21.0    21.8    23.7    27.4;
0.060   20.0    20.8    22.9    26.8;
0.070   19.3    20.1    22.4    26.4;
0.080   18.7    19.6    22.0    26.1;
0.090   18.2    19.1    21.7    25.8;
0.100   17.8    18.8    21.4    25.6;
0.200   15.8    17.0    20.1    24.7;
1.000   14.2    15.5    19.1    24.0;
];

z_vec=table_dev_lam(:,1);
aspect_vec=[1.0 0.5 0.2 0];
fFRe_tab=table_dev_lam(:,2:end);
Dh=4*b*a/(2*(b+a));  % hydraulic diameter [m]

% this is to match solutions below and above turbulent onset
z=max(min(linspace(L/100, L,
100)/2300/Dh,ones(1,100)*1),ones(1,100)*0.001);
[xi,yi] = meshgrid(aspect_vec,z_vec);
fFRe=interp2(xi,yi,fFRe_tab,aspect,z,'linear');
fD_preturb=mean(fFRe*4/2300); % average Darcy friction factor along
the channel

De=Dh*(2/3+11/24*aspect*(2-aspect));
ReDe=2300*De/Dh;
d=linspace(L/100, L, 100);
A=0.0929+1.0161./(d/Dh);
B=-0.2680-0.3193./(d/Dh);
fF=A.*ReDe.^B;
fD_postturb=mean(fF*4);

dfD_turb_fudge=fD_preturb-fD_postturb;


for ind=1:length(Re)
    if Re(ind)<2300
        %
fD=interp1(table_lam(:,1),table_lam(:,2),aspect,'spline')/Re*4;
        z=max(min(linspace(L/100, L,
100)/Re(ind)/Dh,ones(1,100)*1),ones(1,100)*0.001);
```

```
        [xi,yi] = meshgrid(aspect_vec,z_vec);
        fFRe=interp2(xi,yi,fFRe_tab,aspect,z,'linear');
        fD(ind)=mean(fFRe*4/Re(ind)); % average Darcy friction factor
along the channel
        fully_developed_frac(ind)=nnz(z>0.08)/length(z);  % Incropera
pg 72
    else  % and this is the turbulent case
        De=Dh*(2/3+11/24*aspect*(2-aspect));
        ReDe=Re(ind)*De/Dh;
        d=linspace(L/100, L, 100);
        A=0.0929+1.0161./(d/Dh);
        B=-0.2680-0.3193./(d/Dh);
        fF=A.*ReDe.^B;
        fD(ind)=mean(fF*4)+dfD_turb_fudge;
        fully_developed_frac(ind)=nnz(d/Dh>10)/length(d);
    end
end
```

```
function [Nu, fully_developed_frac]=get_Nu(Re, Pr, b, a, L);
% function [Nu, fully_developed_frac]=get_Nu(Re, Pr, b, a, L);
% Calculate Nusselt number for developing flow.
%   (uses Phillips 1988 results, from Incropera)
% Averages the result over the channel length.
%
% outputs:
% Nu=Nusselt number (vector)
% fully_developed_frac=fraction of the length of the channel over
which the
%   flow is considered fully developed
%
% inputs:
% Re=Reynolds number, calculated with hydraulic diameter (vector)
% Pr=Prandtl number
% b=channel height [m]
% a=channel width [m]
% L=channel length [m]

aspect=min([a b])/max([a b]);

% this one is used for the laminar correlations
% z A=1.0 A=0.5 A=0.333 A=0.25 A<0.1
table_dev_lam=[         % pg 82 of Incropera (developing flow), ignore
3-wall effects
0.00010   25.2  23.7    27.0    26.7    31.4;
0.00250   8.9   9.2     9.9     10.4    11.9;
0.00500   7.10  7.46    8.02    8.44    10.0;
0.00556   6.86  7.23    7.76    8.18    9.8;
0.00625   6.60  6.96    7.50    7.92    9.5;
0.00714   6.32  6.68    7.22    7.63    9.3;
```

```
0.00833   6.02  6.37    6.92    7.32    9.1;
0.01000   5.69  6.05    6.57    7.00    8.8;
0.01250   5.33  5.70    6.21    6.63    8.6;
0.01670   4.91  5.28    5.82    6.26    8.5;
0.02500   4.45  4.84    5.39    5.87    8.4;
0.03300   4.18  4.61    5.17    5.77    8.3;
0.05000   3.91  4.38    5.00    5.62    8.25;
0.10000   3.71  4.22    4.85    5.45    8.24;
1.00000   3.60  4.11    4.77    5.35    8.23;
];

z_vec=table_dev_lam(:,1);
aspect_vec=[1.0 0.5 0.333 0.25 0];
Nu_tab=table_dev_lam(:,2:end);
Dh=4*b*a/(2*(b+a));  % hydraulic diameter [m]

% get Nu right before and after turbulence kicks in (for matching
purposes)
z=max(min(linspace(L/100, L,
100)/2300/Dh/Pr,ones(1,100)*1),ones(1,100)*0.0001);
[xi,yi] = meshgrid(aspect_vec,z_vec);
Nu_vec=interp2(xi,yi,Nu_tab,aspect,z,'linear');
Nu_preturb=mean(Nu_vec); % average along the channel
fD_postturb=get_fD(2300, b, a, L);
Nu_postturb=(fD_postturb/8)*(2300-
1000)*Pr*(1+(Dh/L)^(2/3))/(1+12.7*sqrt(fD_postturb/8)*(Pr^(2/3)-1));
delta_Nu_fudge=Nu_postturb-Nu_preturb;  % this is to eliminate an
unphysical discontinuity in the Nu across turbulent regime

for ind=1:length(Re)
    if Re(ind)<2300  % use Phillips table
        z=max(min(linspace(L/100, L,
100)/Re(ind)/Dh/Pr,ones(1,100)*1),ones(1,100)*0.0001);
        [xi,yi] = meshgrid(aspect_vec,z_vec);
        Nu_vec=interp2(xi,yi,Nu_tab,aspect,z,'linear');
        Nu(ind)=mean(Nu_vec); % average along the channel
        fully_developed_frac(ind)=nnz(z>0.05)/length(z);  % Incropera
pg 73
    elseif (Re(ind)>=2300)&(Re(ind)<5e6)  % use modified Gnielinski,
Incropera pg 83
 %        fD=4*(1.58*log(Re(ind))-3.28)^(-2);  % suggested friction
factor
 %       fD=4*(1.82*log(Re(ind))-1.64)^(-2);  % suggested friction
factor
        fD=get_fD(Re(ind), b, a, L);    % the actual Darcy/Moody
friction factor
        d=linspace(L/100, L, 100);
        Nu(ind)=(fD/8)*(Re(ind)-
1000)*Pr*(1+(Dh/L)^(2/3))/(1+12.7*sqrt(fD/8)*(Pr^(2/3)-1))-
delta_Nu_fudge;
        fully_developed_frac(ind)=nnz(d/Dh>10)/length(d);
%    else % for really high Re, use Dittus-Boelter (Incropera, pg 82)
```

```
%        d=linspace(L/100, L, 100);
%        Nu(ind)=0.023*(Re(ind)^(4/5))*Pr^0.4;
%        fully_developed_frac(ind)=nnz(d/Dh>10)/length(d);
    end
end
if Nu<0
    disp('Nu<0 found...  check that Nu correlation applies to chosen
coolant properties.');
end
```

****

Chapter 4 code:

Hi-Z module performance prediction

```
% Hi_Z_module.m
%
% This code estimates the performance of the HZ-2 commercial generator
from
% measurements on its constituent n-type and p-type elements, and from
% measured geometrical parameters.  The data files are not included.

% all units are in cm, K, Volts, Watts, etc.


AN=0.145^2; % area of an n-type leg
AP=0.145^2; % area of a p-type leg
Amod=(2.9^2);  % area of the module
gap=0.068;  % distance
LN=0.281; % length of n-type leg
LP=0.281;  % length of p-type leg
Lcontact=0.077; % thickness of metal contacts
Lmod=0.457;  % module length (element + contacts)
Aspread=(0.15+gap)^2; % the area above one element in the checker-
board
kappa_kapton=0.006;  % book value for thermal conductivity of the egg-
crate
A_kapton=2.9^2-(AN+AP)*97;  % total are of the Kapton
RthN_parasitic=1.5/AN*2; % thermal parasitic contact, due to indium
foil
RthP_parasitic=1.5/AP*2; % thermal parasitic contact, due to indium
foil
Rthmod_parasitic=0.32; % estimated from Hi-Z website ~ 10/200 degree
drop
```

```
Qpar=3.91*0.3167/6*0.2;  % measured parasitic shunt thermal conduction
                         %   between contacts


dataN=load('HZn');
% temperature data for the n-type element
ThN=dataN(:,1); TcN=dataN(:,2); dTN=ThN-TcN; T1N=dataN(:,3);
T2N=dataN(:,5);
OCVN=dataN(:,6);  % open circuit voltage
VN=dataN(:,8);  % closed circuit voltage
IN=dataN(:,9);  % closed circuit current
PgenN=VN.*IN;  % generated power
loadN=VN./IN;  % milli-ohm load resistance
time=dataN(:,12);  % time stamp
rN=235:246;  % thermally stable time (used for quantitative
measurements)
dTN_max=mean(dTN(rN)); % temperature difference
VN_max=mean(VN(rN));  % closed circuit voltage
loadN_max=mean(loadN(rN));  % load resistance
OCVN_max=mean(OCVN(rN));    % open circuit voltage
fN=VN_max/OCVN_max;    % measure of impedance mismatch
RintN=loadN_max*(1-fN)/fN;  % impedance matched internal load
                            %  (includes lead resistance)
SeebeckN_max=mean(OCVN(rN)./dTN(rN)) % Seebeck estimate
conductivityN_max=1/((RintN)/LN*AN)  % conductivity estimate
PgenN_max=mean(PgenN(rN))  % power generated
QtotoutN=3.91*0.3167/6*mean(T1N(rN)-T2N(rN));  % measured heat flow
QoutN=QtotoutN-Qpar;  % true heat flow
EffN_max=PgenN_max/QoutN  % efficiency estimate
kappaN=QoutN*LN/AN/dTN_max  % thermal conductivity estimate (w/
interface)
Rth_N=LN/AN/kappaN;  % thermal impedance of element + interface
dTN_nothpar=dTN_max*Rth_N/(Rth_N+RthN_parasitic);
kappaN_nothpar=QoutN*LN/AN/dTN_nothpar  % thermal cond w/o interface

dataP=load('HZp3');
% temperature data for p-type element; see n-type documentation
ThP=dataP(:,1);
TcP=dataP(:,2);
dTP=ThP-TcP;
T1P=dataP(:,3);
T2P=dataP(:,5);
OCVP=dataP(:,6);
VP=dataP(:,8);
IP=dataP(:,9);
PgenP=VP.*IP;
loadP=VP./IP;
time=dataP(:,12);
rP=480:560;
dTP_max=mean(dTP(rP));
```

```
VP_max=mean(VP(rP));
loadP_max=mean(loadP(rP));
OCVP_max=mean(OCVP(rP));
fP=VP_max/OCVP_max;
RintP=loadP_max*(1-fP)/fP;
SeebeckP_max=mean(OCVP(rP)./dTP(rP))
conductivityP_max=1/((RintP)/LP*AP)
PgenP_max=mean(PgenP(rP))
QtotoutP=3.91*0.3167/6*mean(T1P(rP)-T2P(rP));
QoutP=QtotoutP-Qpar;
kappaP=QoutP*LP/AP/dTP_max
Rth_P=LP/AP/kappaP;
EffP_max=PgenP_max/QoutP
dTP_nothpar=dTP_max*Rth_P/(Rth_P+RthP_parasitic);
kappaP_nothpar=QoutP*LP/AP/dTP_nothpar

% series spreading resistance from each element
Rth_spread=2*getRspread(AN,Aspread,Lcontact,0.8);

% parallel thermal conduction through kapton
Rth_kapton=LN/A_kapton/kappa_kapton;

% total parallel thermal resistance of the elements (excluding
parasitics)
Rth_mod_nopar=1/(97/(Rth_N+Rth_spread)+97/(Rth_P+Rth_spread)+1/Rth_kap
ton);

% total parallel thermal resistance of the elements (including
parasitics)
Rth_mod=1/(97/(Rth_N+Rth_spread)+97/(Rth_P+Rth_spread)+1/Rth_kapton)+R
thmod_parasitic


%Rth_mod_meas=2.78;
kappa_mod=Lmod/Rth_mod/Amod  % module thermal conductivity

% estimate of the importance of thermal interface resistance in
calculation
frac_dT_mod=Rth_mod_nopar/Rth_mod

V_mod=(-VP_max+VN_max)*97*frac_dT_mod
OCV_mod=(-OCVP_max+OCVN_max)*97*frac_dT_mod
Pgen_mod=(PgenP_max+PgenN_max)*97*(frac_dT_mod)^2
load_mod=(RintP+RintN)*97
load_mod_nopar=(RintP-9e-3+RintN-9e-3)*97
Qout_mod=200/Rth_mod
Eff_mod=Pgen_mod/Qout_mod*100

% these are the book values for BiTe
SbiteP=230e-6;
condbiteP=1/1.75e-3;
kbiteP=0.012;
```

```
SbiteN=-200e-6;
condbiteN=1/1.35e-3;
kbiteN=0.014;

% these are the datasheet values for module performance
Pgen_book=2.5;
Resist_book=4.0;
OCV_book=6.53;
Eff_book=4.5;


figure(10);
bar_mat=[-SeebeckP_max/SbiteP conductivityP_max/condbiteP
kappaP/kbiteP ; -SeebeckN_max/SbiteN conductivityN_max/condbiteN
kappaN/kbiteN];
bar(bar_mat);


figure(11);
bar_mat2=[Pgen_mod/Pgen_book Eff_mod/Eff_book load_mod/Resist_book
OCV_mod/OCV_book];
bar(bar_mat2);
```

****

Chapter 5 code:

1) Function predicting errors of thermoreflectance measurements

2) Utility that is used to graphically select a region of a figure and return the correct |ΔR/R|

3) Function that corrects for the finite shutter time in the camera (~5% correction)

4) Procedure used for modeling of the dynamic temperature distribution in MAM material

5) 1-D finite difference time domain Crank-Nicolson solver

6) 1-D finite difference time domain Crank-Nicolson solver with arbitrary time and space step-sizes

---

Function predicting errors of thermoreflectance measurements

```
function [oerr_pct rerr_pct]=predict_error(dR_R_in, c_in, N_in, alpha,
beta)
% function [oerr_pct rerr_pct]=predict_error(dR_R_in, c_in, N_in,
alpha, beta)
%
```

```
% This function plots the error bars of the thermoreflectance
measurement
%   given a light level, signal level, and noise parameters.  It is
%     vectorized.
%
% outputs:
%   oerr_pct - the predicted offset error (as pct of dR/R)
%   rerr_pct - the standard deviation of the random error (as pct of
dR/R)
%
% inputs:
%   dR_R_in - expected signal level
%   c_in - dc light level (for 12-bit camera, max of 4096)
%   N_in - vector with the numbers iterations to check
%   alpha, beta - camera noise parameters (defined in JOSA A paper)


% use Dietrich's noise numbers
if nargin<4
    A=0.1063;
    B=3.936;
    alpha=A*4/pi/sqrt(2-pi/2); % correct for difference in variance of
DRR versus variance of fundamental noise
    beta=B*4/pi/sqrt(2-pi/2);
end
if nargin<3
    N_in=logspace(0,log10(200000),100);
end
if nargin<3
    c_in=4096;
end

warning off;

% calculate the symbolic mean and standard deviation of the
%   relevant probability distribution
syms m x s;
Rician=x/s^2*besseli(0,m*x/s^2)*exp(-(x^2+m^2)/2/s^2);
exp1=Rician*x;
moment1_Ric=int(exp1,'x',0,inf);
exp2=Rician*x^2;
moment2_Ric=int(exp2,'x',0,inf);

% need an expression for the first moment which doesn't mess up MATLAB
for
% large input arguments

z=(m/s/2)^2;
I_asym='exp(z)/sqrt(2*pi*z)*(1-(mu-1)/8/z+(mu-1)*(mu-9)/2/(8*z)^2-(mu-
1)*(mu-9)*(mu-25)/6/(8*z)^3)';
I_asym=subs(I_asym,z);
mu0=0;
```

```
mu1=4;
I0_asym=subs(I_asym,'mu',mu0);
I1_asym=subs(I_asym,'mu',mu1);
moment1_Ric_asym='1/4/s*exp(-
1/4/s^2*m^2)*2^(1/2)*pi^(1/2)*((2*s^2+m^2)*I0_asym+m^2*I1_asym)';
moment1_Ric_asym=subs(moment1_Ric_asym,'I0_asym',I0_asym);
moment1_Ric_asym=subs(moment1_Ric_asym,'I1_asym',I1_asym);
%moment1_Ric_asym=simple(moment1_Ric_asym);
%pause(1);
moment1_Ric_asym='s/16*(8*s/m+2*(s/m)^3+3*(s/m)^5+75*(s/m)^7+16*(m/s))
';

% this is just the zero TR signal case (just noise)
moment1_Ray=limit(moment1_Ric,'m',0);
moment2_Ray=limit(moment2_Ric,'m',0);


dR_R=dR_R_in;
N=N_in;
c=c_in;

for iN=1:length(N)

% for a given (real) TR signal, what is the spurious DC component and
the
% variance?  Calculate this for a family of possible TR magnitudes...
            N_I=N(iN); %N_I=logspace(1,7,8);%2*10^4;  % number of
iterations
            mean_I=c.*N_I;  % mean_I=(I1+I2+I3+I4)/4
            s_CCD=(alpha*sqrt(c)+beta);  % standard deviation of CCD
noise (units of CCD counts)
            v_CCD=s_CCD^2;
            v_I=v_CCD.*N_I;  % variance in I1, I2, ...
            v_2I=v_I*2;  % variance in I0-I2, I1-I3
            dI=4/pi*dR_R.*mean_I; % dI=sqrt(((I0-I2)^2+(I1-I3)^2)/2)
            denom_const=sqrt(2)/pi*mean_I*4;
            norm_v_2I=v_2I./denom_const.^2;
            m_val=dR_R;
            s_val=sqrt(norm_v_2I);

            m=m_val;
            s=s_val;
            if ((m/s)^2)>40  % use asymptotic expansion for first
moment
                meas_dR_R=eval(moment1_Ric_asym);
            else
                meas_dR_R=eval(moment1_Ric);  % use normal expression
for first moment
            end
            mom2=2*s^2+m^2;
            std_dR_R=sqrt((mom2-meas_dR_R.^2));
            oerr(iN)=(meas_dR_R-dR_R);  % offset error
```

```
            rerr(iN)=std_dR_R;    % random error
end

        oerr_pct=oerr/dR_R*100;
        rerr_pct=rerr/dR_R*100;

        figure;
        h=loglog(N,dR_R+rerr,'g:',N,dR_R-
rerr,'g:',N,dR_R+oerr,'k',N,dR_R*ones(1,length(N)),'b-');
        xlabel('N');
        ylabel('\DeltaR/R');
        title(['CCD counts = ' num2str(c) ', \DeltaR/R = '
num2str(dR_R)]);
        set(gca,'YScale','log');
        set(gca,'XScale','log');
        set(h,'LineWidth',2);
        set(gca,'LineWidth',2);
        set(gca,'FontSize',14);

        figure;
        h=semilogx(N,oerr_pct,'k',N,rerr_pct,'g');
        xlabel('N');
        ylabel('Offset and Random Error (% of \DeltaR/R)');
        set(h,'LineWidth',2);
        set(gca,'LineWidth',2);
        set(gca,'FontSize',14);
        title(['CCD counts = ' num2str(c) ', \DeltaR/R = '
num2str(dR_R)]);
        set(gca,'YScale','linear');
        set(gca,'XScale','log');
        set(h,'LineWidth',2);
        set(gca,'LineWidth',2);
        set(gca,'FontSize',14);

warning on;
```

Utility that is used to graphically select a region of a figure and return the correct |ΔR/R|

```
% This utility allows me to graphically select a region of the
currently
%       loaded figure and get the correctly pre-averaged
thermoreflectance
% It works only if the plot is vs. indices (not real length vectors)
%  Also outputs the "naive" (high-offset error) post-averaged TR.

figure(10);  % this figure is some thing with the same size as I1, I2,
etc,
%               and whose x and y coords are the matrix indices)
```

```
k = waitforbuttonpress;
point1 = get(gca,'CurrentPoint');     % button down detected
finalRect = rbbox;                    % return figure units
point2 = get(gca,'CurrentPoint');     % button up detected
point1 = point1(1,1:2);               % extract x and y
point2 = point2(1,1:2);
p1 = min(point1,point2);              % calculate locations
offset = abs(point1-point2);          % and dimensions
x = [p1(1) p1(1)+offset(1) p1(1)+offset(1) p1(1) p1(1)];
y = [p1(2) p1(2) p1(2)+offset(2) p1(2)+offset(2) p1(2)];
%axis([p1(1) p1(1)+offset(1) p1(2) p1(2)+offset(2)]);

hold on
axis manual
plot(x,y,'w:');

xindmin=ceil(p1(1));
xindmax=floor(p1(1)+offset(1));
yindmin=ceil(p1(2));
yindmax=floor(p1(2)+offset(2));

disp('Finding mean thermoreflectance of portion of image using ACDC,
I1, I2, I3, I4.')
disp('Make sure data is plotted versus matrix indices (not real
distance)!');
fprintf(1,'\n');

mean_from_ACDC=mean(mean(ACDC(yindmin:yindmax,xindmin:xindmax)));
%fprintf(1,'mean from ACDC average: %.5e \n',mean_from_ACDC);

mI=zeros(1,4);
mI(1)=mean(mean(I1(yindmin:yindmax,xindmin:xindmax)));
mI(2)=mean(mean(I2(yindmin:yindmax,xindmin:xindmax)));
mI(3)=mean(mean(I3(yindmin:yindmax,xindmin:xindmax)));
mI(4)=mean(mean(I4(yindmin:yindmax,xindmin:xindmax)));

mean_from_I1toI4=pi/sqrt(2)*sqrt((mI(3)-mI(1))^2+(mI(4)-
mI(2))^2)/sum(mI);
%fprintf(1,'mean from I1, I2, I3, I4 average: %.5e
\n',mean_from_I1toI4);

if exist('corr_fact')
    corr_mean_ACDC=mean_from_ACDC*corr_fact;
    corr_mean_I1toI4=mean_from_I1toI4*corr_fact;
    fprintf(1,'corrected mean from TR (naive) average: %.5e
\n',corr_mean_ACDC);
    fprintf(1,'corrected mean from I1, I2, I3, I4 average: %.5e
\n',corr_mean_I1toI4);

end
```

```
hold off;
```

Function that corrects for the finite shutter time in the camera (~5% correction)

```
function
corr_factor=finite_shutter_correction(camera_trig_freq,shutter_open_ti
me)

%
% This simulates the effects of fractional aquisition on the Opteon
lock-in TR measurement,
%   meaning that the camera shutter is only open for a small time out
of the total measurement.
%
% This is useful for slow measurements (e.g. 2 Hz camera trigger), and
also for faster
%   measurements if you want to play it safe with the camera and avoid
missed triggers.
%
% The real output of the code is a correction factor which you simply
scale your measures DR/R by
%   to obtain the true DR/R.
%
Ttrig=1/camera_trig_freq;
w=camera_trig_freq/4*2*pi;
samplet=linspace(0,3*Ttrig,4);  % times that camera is triggered
(opening of shutter)
exposure=shutter_open_time;  % exposure time [s]
exposuret=samplet+exposure;  % ends of the time shutter is open

%real_TR=max(real(dT.*beta))/R0  % defined as amplitude, not pk-to-
pk...

% analytic calculation of (easy sine function) integrals (the
numerical calculation just checks this)
DC=1; % this represents the magnitude of the DC light on the CCD
AC=1e-5; % this represents the amplitude of the AC light on the CCD

% trivial integral worked out in lab notebook
real_TR=AC/DC;
for ind=1:4
   intsig_theory(ind)=DC*exposure+AC/w*(cos(w*samplet(ind))-
cos(w*(exposuret(ind))));   % measure of accumulated CCD counts
end
TRmeas_analytic=pi/sqrt(2)*sqrt((intsig_theory(1)-
intsig_theory(3))^2+(intsig_theory(2)-
intsig_theory(4))^2)/(intsig_theory(1)+intsig_theory(2)+intsig_theory(
3)+intsig_theory(4));
```

```
corr_factor=real_TR/TRmeas_analytic;
disp(['Measured signal should be multiplied by
',num2str(corr_factor),' to get the true dR/R.']);
```

Procedure used for modeling of the dynamic temperature distribution in MAM material

---

```
% this code calls the 1-D finite difference time domain solver in
order to
% simulate the dymnamic response of the Bi2Te3 MAM element to an
applied
% current.

clear all;
f=0.5;
tmax=1/f*12;
nt=1024;
t=linspace(0,tmax,nt);
nx=100;
J=200*sin(2*pi*f*t+pi/4);
%J=500*[-1*ones(1,floor(nt/2)) ones(1,ceil(nt/2))];
%J=-500*[ones(1,nt)];

L_BiTe=0.0508;
dx=L_BiTe/(nx-1);
alpha_BiTe=240e-6;
K_BiTe=0.0147;
sigma_BiTe=737;
cprho_BiTe=K_BiTe/0.011;
Tamb=300;
h=10; hleft=h*1.5; hright=h;
C=0.385*8.96*0.03;
Cleft=C; Cright=C;

%[Rpleft, Cpleft]=get_convertedRC(1/hleft,Cleft,f);
%[Rpright, Cpright]=get_convertedRC(1/hright,Cright,f);
%hpright=1/Rpright;
%hpleft=1/Rpleft;

TE_time_const=cprho_BiTe*(L_BiTe^2)/K_BiTe
bc_time_const=C/h
excitation_time_const=1/f
Rc=0;%5e-5;
%hc=1000;

mat_struct=[
%    2*dx alpha_BiTe hc*dx*2 1/(Rc/dx) cprho_BiTe;
%    L_BiTe-2*L_BiTe/nx*2 alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
```

```
%    2*dx alpha_BiTe hc*dx*2 1/(Rc/dx) cprho_BiTe;
%    2*dx alpha_BiTe K_BiTe 1/(Rc/dx) cprho_BiTe;
%    L_BiTe-2*L_BiTe/nx*2 alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
    L_BiTe alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
%    2*dx alpha_BiTe K_BiTe 1/(Rc/dx) cprho_BiTe;
];

bcvec=[3 Tamb Tamb hleft hright Cleft Cright];

[T,x,t]=CNsolver(nt,nx,mat_struct,J,bcvec,tmax);

Tinit=T(:,end);
clear T;

[T,x,t]=CNsolver(nt,nx,mat_struct,J,bcvec,tmax,Tinit);

fs=1/(tmax/nt);
fplot=linspace(0,fs,nt);

f1_index=f/(fs/nt)+1;
f2_index=2*f/(fs/nt)+1;

FFT_mag=abs(fft(T',nt))*2/nt;
figure(2);
subplot(2,2,1);
hold on;
h11=plot(x*1e4,FFT_mag(f1_index,:)','b');
%v=axis;
%axis([0 508 0 2]);
hold off;

subplot(2,2,2);
hold on;
h12=plot(x*1e4,FFT_mag(f2_index,:)','b');
%v=axis;
%axis([0 508 0 1]);
hold off;

c1=220;
FFT_phase=unwrap(angle(fft(T',nt)))*180/pi;
subplot(2,2,3);
hold on;
h21=plot(x*1e4,FFT_phase(f1_index,:),'b');
%axis([0 508 -95 95]);
hold off;

subplot(2,2,4);
hold on;
h22=plot(x*1e4,FFT_phase(f2_index,:),'b');
v=axis;
%axis([0 508 -90 90]);
hold off;
```

```
clear all;
f=7;
tmax=1/f*12;
nt=1024;
t=linspace(0,tmax,nt);
nx=100;
J=200*sin(2*pi*f*t+pi/4);
%J=500*[-1*ones(1,floor(nt/2)) ones(1,ceil(nt/2))];
%J=-500*[ones(1,nt)];

L_BiTe=0.0508;
dx=L_BiTe/(nx-1);
alpha_BiTe=240e-6;
K_BiTe=0.0147;
sigma_BiTe=737;
cprho_BiTe=K_BiTe/0.011;
Tamb=300;
h=10; hleft=h*1.5; hright=h;
C=0.385*8.96*0.03;
%C=0.066*0.5;
Cleft=C; Cright=C;

%[Rpleft, Cpleft]=get_convertedRC(1/hleft,Cleft,f);
%[Rpright, Cpright]=get_convertedRC(1/hright,Cright,f);
%hpright=1/Rpright;
%hpleft=1/Rpleft;

TE_time_const=cprho_BiTe*(L_BiTe^2)/K_BiTe
bc_time_const=C/h
excitation_time_const=1/f
Rc=0;%5e-5;
%hc=1000;

mat_struct=[
%    2*dx alpha_BiTe hc*dx*2 1/(Rc/dx) cprho_BiTe;
%    L_BiTe-2*L_BiTe/nx*2 alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
%    2*dx alpha_BiTe hc*dx*2 1/(Rc/dx) cprho_BiTe;
%    2*dx alpha_BiTe K_BiTe 1/(Rc/dx) cprho_BiTe;
%    L_BiTe-2*L_BiTe/nx*2 alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
    L_BiTe alpha_BiTe K_BiTe sigma_BiTe cprho_BiTe;
%    2*dx alpha_BiTe K_BiTe 1/(Rc/dx) cprho_BiTe;
];

bcvec=[3 Tamb Tamb hleft hright Cleft Cright];

[T,x,t]=CNsolver(nt,nx,mat_struct,J,bcvec,tmax);

Tinit=T(:,end);
clear T;
```

```
[T,x,t]=CNsolver(nt,nx,mat_struct,J,bcvec,tmax,Tinit);

fs=1/(tmax/nt);
fplot=linspace(0,fs,nt);

f1_index=f/(fs/nt)+1;
f2_index=2*f/(fs/nt)+1;

FFT_mag=abs(fft(T',nt))*2/nt;
figure(2);
subplot(2,2,1);
hold on;
h11=plot(x*1e4,FFT_mag(f1_index,:)','b');
%v=axis;
%axis([0 508 0 2]);
hold off;

subplot(2,2,2);
hold on;
h12=plot(x*1e4,FFT_mag(f2_index,:)','b');
%v=axis;
%axis([0 508 0 1]);
hold off;

c1=220;
FFT_phase=unwrap(angle(fft(T',nt)))*180/pi;
subplot(2,2,3);
hold on;
h21=plot(x*1e4,FFT_phase(f1_index,:),'b');
%axis([0 508 -95 95]);
hold off;

subplot(2,2,4);
hold on;
h22=plot(x*1e4,FFT_phase(f2_index,:),'b');
v=axis;
%axis([0 508 -90 90]);
hold off;
```

1-D finite difference time domain Crank-Nicolson solver

```
function [T,x,t]=CNsolver(nt,nx,sample_struct,J,bc,tmax,Tinit,Jinit)
% [T,x,t]=CNsolver(nt,nx,sample_struct,J,bc,tmax,Tinit,Jinit)
%
% This function solves the following equation for heat and electrical
% transport in a TE element in 1D:
% density*cp*dT/dt = d/dx(K*dT/dx)+J^2*rho-J*T*d(alpha)/dx
%
```

```
%  This is useful for predicting the temperature distribution given a
bias
%  current density J, and boundary conditions at the ends of the
sample.
%  All of the material parameters can spatially vary.
%
%  T - matrix of temperatures, T(space index, time index)
%  x - output x (position) vector
%  t - output t (time) vector
%
%  nt - number of time steps
%  nx - number of x steps (this will later be changed to allow for
%   nonuniform grids)
%  sample_struct - this describes the material parameters, and has the
structure:
%       [L1 alpha1 K1 sigma1 C; L2 alpha2 K2 sigma2 ; ...]
%            Here alpha=Seebeck coeff [V/K], sigma=conductivity
[1/ohm/cm],
%            K=thermal conductivity [W/cm/K], C=density*cp=heat
capacity
%            per unit volume [J/cm^3/K]
%  J = assumed current density, vector of length nt [A/cm^2]
%  bc - boundary conditions; bc = [{1,2,3} bcparams]
%       1 - constant temperature (Dirichlet); bcparams = [T_0 T_N]
%       2 - heat sink (convective, Neumann-type b.c.); bcparams = [T_0
T_N h_0 h_N]
%       3 - addl volumetric heat cap.; bcparams = [T_0 T_N h_0 h_N c_0
c_N]
%            here T [W], h [W/cm^2/K],
%            C=cp*density*volume/(power flux in area) [J/cm^2/K]
%  tmax - time to integrate the equations up to
%  Tinit - initial temperature for the material in K (vector of length
nx)
%  Jinit - initial current (taken to be at -dt)
%
%  modifications:
%   7/31/06: changed b.c's to include Ohmic heating and satisfy C-N
condition
%       note that for Dirichlet conditions,
%

if nargin<7
    Tinit=ones(nx,1)*300;
end

if nargin<8
    Jinit=0;
end


L=sum(sample_struct(:,1));
dx=L/(nx-1);
```

```
dt=tmax/(nt-1);

x=linspace(0,L,nx)';
t=linspace(-dt,tmax,nt+1);
T=zeros(nx,nt+1);

% now scan through the structure file and build the material x-profile
alpha=zeros(nx,1); K=zeros(nx,1); rho=zeros(nx,1); C=zeros(nx,1);
bkpoint=0; pos=1;
for ind=1:size(sample_struct,1)
    bkpoint=bkpoint+sample_struct(ind,1);
    while ((pos-1)*dx)<=bkpoint+1e-12 % the 1e-12 is there to take
care of fp inaccuracies
        alpha(pos)=sample_struct(ind,2);
        K(pos)=sample_struct(ind,3);
        rho(pos)=1/sample_struct(ind,4);
        C(pos)=sample_struct(ind,5);
        pos=pos+1;
    end
end

% get the half step parameters
alpha_hs=zeros(nx-1,1); K_hs=zeros(nx-1,1);
rho_hs=zeros(nx-1,1); C_hs=zeros(nx-1,1);
for ind=1:(nx-1)
    alpha_hs(ind)=(alpha(ind)+alpha(ind+1))/2;
    K_hs(ind)=(K(ind)+K(ind+1))/2;
    rho_hs(ind)=(rho(ind)+rho(ind+1))/2;
    C_hs(ind)=(C(ind)+C(ind+1))/2;
end

T(:,1)=Tinit;    % set t=0 boundary condition

% to prevent oscillations, start at zero (at -dt) and add low pass
%       filtered (half-step) currents
J_hs=[Jinit filter([1/2 1/2],1,J)];


for ind=2:nt+1
    % build tridiagonal X matrix for interior points of material
    X=zeros(nx,nx);
    Xa=diag([K_hs(1:end-1) ; 0]/2,-1);  % so-called a terms
(prediagonal)
    Xc=diag([0 ; K_hs(2:end)/2],1);  % so-called c terms
(postdiagonal)
    % and now the b terms (diagonal)
    bvec=(1/2)*(-K_hs(1:end-1)-K_hs(2:end))-...
        J_hs(ind)/2*dx*(alpha_hs(2:end)-alpha_hs(1:end-1))-C(2:end-
1)*dx^2/dt;
    Xb=diag([0 ; bvec ; 0]);
    X=Xa+Xb+Xc;
    % now set the boundary conditions in the X matrix...
```

```
    T1=bc(2); TN=bc(3); % (all bc types have temperatures)

    switch bc(1)
     case 1  % simple Dirichlet
         b1=1*dx; c1=0*dx;
         bN=1*dx; aN=0*dx;
     case 2  % simple Neumann
         h1=bc(4); hN=bc(5);

b1=C(1)*dx/2/dt*dx+J_hs(ind)/2*alpha_hs(1)*dx+K_hs(1)/2+h1/2*dx;
         c1=-K_hs(1)/2;
         bN=C(end)*dx/2/dt*dx-
J_hs(ind)/2*alpha_hs(end)*dx+K_hs(end)/2+hN/2*dx;
         aN=-K_hs(end)/2;
     case 3  % Neumann w/ heat capacitance
         h1=bc(4); hN=bc(5); C1=bc(6); CN=bc(7);

b1=(C(1)*dx/2+C1)/dt*dx+J_hs(ind)/2*alpha_hs(1)*dx+K_hs(1)/2+h1/2*dx;
         c1=-K_hs(1)/2;
         bN=(C(end)*dx/2+CN)/dt*dx-
J_hs(ind)/2*alpha_hs(end)*dx+K_hs(end)/2+hN/2*dx;
         aN=-K_hs(end)/2;
    end
    X(1,1)=b1; X(1,2)=c1; X(end, end-1)=aN; X(end, end)=bN;



    % build the d-vector at interior points
    d=zeros(nx,1);
    d(2:end-1)=-J_hs(ind)^2*dx^2*rho(2:end-1)/2-J_hs(ind-
1)^2*dx^2*rho(2:end-1)/2-...
        T(3:end,ind-1).*K_hs(2:end)/2-...
        T(1:end-2,ind-1).*K_hs(1:end-1)/2+T(2:end-1,ind-
1).*((K_hs(1:end-1)+...
        K_hs(2:end))/2+J_hs(ind-1)*(alpha_hs(2:end)-alpha_hs(1:end-
1))/2*dx-...
        C(2:end-1)/dt*dx^2);

    % now set the boundary conditions in the d vector...
    switch bc(1)
        case 1  % simple Dirichlet
            d1=T1*dx; dN=TN*dx;
            %  the next two lines add a little heating for the last
dx's
            d(2)=d(2)-(J_hs(ind)^2+J_hs(ind-1)^2)/2*dx^2*rho(end)/2;
            d(end-1)=d(end-1)-(J_hs(ind)^2+J_hs(ind-
1)^2)/2*dx^2*rho(end)/2;
        case 2  % simple Neumann
            d1=T(1,ind-1)*C(1)*dx/2/dt*dx+h1*T1*dx-h1/2*T(1,ind-
1)*dx+...
                K_hs(1)/2*(T(2,ind-1)-T(1,ind-1))-...
                J_hs(ind-1)/2*alpha_hs(1)*T(1,ind-1)*dx+...
```

```
                    (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx^2*rho(1)/2;
                dN=T(end,ind-1)*C(end)*dx/2/dt*dx+hN*TN*dx-hN/2*T(end,ind-
1)*dx+...
                    K_hs(end)/2*(T(end-1,ind-1)-T(end,ind-1))+...
                    J_hs(ind-1)/2*alpha_hs(end)*T(end,ind-1)*dx+...
                    (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx^2*rho(end)/2;
        case 3  % Neumann w/ heat capacitance
                d1=T(1,ind-1)*(C1+C(1)*dx/2)/dt*dx+h1*T1*dx-h1/2*T(1,ind-
1)*dx+...
                    K_hs(1)/2*(T(2,ind-1)-T(1,ind-1))-...
                    J_hs(ind-1)/2*alpha_hs(1)*T(1,ind-1)*dx+...
                    (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx^2*rho(1)/2;
                dN=T(end,ind-1)*(CN+C(end)*dx/2)/dt*dx+hN*TN*dx-
hN/2*T(end,ind-1)*dx+...
                    K_hs(end)/2*(T(end-1,ind-1)-T(end,ind-1))+...
                    J_hs(ind-1)/2*alpha_hs(end)*T(end,ind-1)*dx+...
                    (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx^2*rho(end)/2;
    end
    d(1)=d1; d(end)=dN;
 %   keyboard;
    [L,U]=lu(X);  % gets upper and lower triangular factorization
    T(:,ind)=U\(L\d);  % calculate current temperature from previous
one
end

% strip off -dt time point used to set up initial conditions
t=t(2:end);
T=T(:,2:end);
```

1-D finite difference time domain Crank-Nicolson solver with arbitrary time and space step-sizes

```
function [T,x,t]=CNsolver(sample_struct,t,J,bc,Tinit,Jinit)
% [T,x,t]=CNsolver(nt,nx,sample_struct,J,bc,tmax,Tinit,Jinit)
%
%  This function solves the following equation for heat and electrical
%  transport in a TE element in 1D:
%  density*cp*dT/dt = d/dx(K*dT/dx)+J^2*rho-J*T*d(alpha)/dx
%
%  This is useful for predicting the temperature distribution given a
bias
%  current density J, and boundary conditions at the ends of the
sample.
%  All of the material parameters can spatially vary.
%
%  T - matrix of temperatures, T(space index, time index)
%  x - output x (position) vector
%  t - output t (time) vector
%
```

```
%  sample_struct - this describes the material parameters, and has the
structure:
%       [n1 L1 alpha1 K1 sigma1 C; n2 L2 alpha2 K2 sigma2 ; ...]
%            Here alpha=Seebeck coeff [V/K], sigma=conductivity
[1/ohm/cm],
%            K=thermal conductivity [W/cm/K], C=density*cp=heat
capacity
%            per unit volume [J/cm^3/K], L= length [cm], n=number of
mesh
%            points
%  t - input t (time) vector
%  J = assumed current density, vector of length nt (matches up with
t) [A/cm^2]
%  bc - boundary conditions; bc = [{1,2,3} bcparams]
%       1 - constant temperature (Dirichlet); bcparams = [T_0 T_N]
%       2 - heat sink (convective, Neumann-type b.c.); bcparams = [T_0
T_N h_0 h_N]
%       3 - addl volumetric heat cap.; bcparams = [T_0 T_N h_0 h_N c_0
c_N]
%            here T [W], h [W/cm^2/K],
%            C=cp*density*volume/(power flux in area) [J/cm^2/K]
%  Tinit - initial temperature for the material in K (vector of length
nx)
%  Jinit - initial current (taken to be at -dt)
%
%  modifications:
%   7/31/06: changed b.c's to include Ohmic heating and satisfy C-N
condition
%       note that for Dirichlet conditions,
%
%   8/02/06: changed b.c.'s to include heat capacity of end elements
and
%       added arbitrary time vector capability

if nargin<6
    Jinit=0;
end

if nargin<5
    Tinit=ones(sum(sample_struct(:,1))+1,1)*300;
end

%------------------------------------
%  This section sets up the structure
%------------------------------------

nx=sum(sample_struct(:,1))+1;
L=sum(sample_struct(:,2));
x=[];
dx=[];
L_pos=0;
```

```
% the half step alphas are the natural ones (each is associated with a
diff
% vol elt)
alpha_hs=[];
K_hs=[];
rho_hs=[];
C_hs=[];

% x is positioned at the ends of every differential volume element

for ind=1:size(sample_struct,1)
    L_seg=sample_struct(ind,2);
    n_seg=sample_struct(ind,1);
    x=[x ; linspace(L_pos,L_pos+L_seg*(1-1/n_seg),n_seg)'];
    L_pos=L_pos+L_seg;

    dx=[dx ; L_seg/n_seg*ones(n_seg,1)];
    alpha_hs=[alpha_hs ; sample_struct(ind,3)*ones(n_seg,1)];
    K_hs=[K_hs ; sample_struct(ind,4)*ones(n_seg,1)];
    rho_hs=[rho_hs ; 1/sample_struct(ind,5)*ones(n_seg,1)];
    C_hs=[C_hs ; sample_struct(ind,6)*ones(n_seg,1)];
end

% since the x mesh can fall on boundaries, these are (sometimes)
averaged
alpha=([alpha_hs ; alpha_hs(end)]+[alpha_hs(1) ; alpha_hs])/2;
K=([K_hs ; K_hs(end)]+[K_hs(1) ; K_hs])/2;
rho=([rho_hs ; rho_hs(end)]+[rho_hs(1) ; rho_hs])/2;
C=([C_hs ; C_hs(end)]+[C_hs(1) ; C_hs])/2;

x=[x ; L];  % add that final x mesh point
nx=length(x);
dx=diff(x);

%-------------------------------------
%keyboard;

% add artificial time points at begining
nt=length(t);
t=[t(1)-(t(2)-t(1)) t t(end)+(t(end)-t(end-1))];
dt=diff(t);
T=zeros(nx,nt+1);

T(:,1)=Tinit;     % set t=0 boundary condition

% to prevent oscillations, start at zero (at -dt) and add low pass
%       filtered (half-step) currents.  (This is a bit of voodoo...)
J_hs=[Jinit filter([1/2 1/2],1,J)];

% useful diff vol elts
dx1=dx(1:end-1);
dx2=dx(2:end);
```

```
dx12=(dx1+dx2)/2;


for ind=2:nt+1


    % build tridiagonal X matrix for interior points of material
    X=zeros(nx,nx);
    Xa=diag([K_hs(1:end-1).*dx12./dx1/2 ; 0],-1);  % so-called a terms
(prediagonal)
    Xc=diag([0 ; K_hs(2:end).*dx12./dx2/2],1);  % so-called c terms
(postdiagonal)
    % and now the b terms (diagonal)
    bvec=-1/2.*dx12.*(K_hs(1:end-1)./dx1+K_hs(2:end)./dx2)-...
        J_hs(ind)*(alpha_hs(2:end)-alpha_hs(1:end-1)).*dx12/2-C(2:end-
1)./dt(ind-1).*(dx12.^2);
    Xb=diag([0 ; bvec ; 0]);
    X=Xa+Xb+Xc;
    % now set the boundary conditions in the X matrix...
    T1=bc(2); TN=bc(3); % (all bc types have temperatures)

    switch bc(1)
     case 1  % simple Dirichlet
         b1=1*dx(1); c1=0*dx(1);
         bN=1*dx(end); aN=0*dx(end);
     case 2  % simple Neumann
         h1=bc(4); hN=bc(5);
         b1=C(1)*dx(1)/2./dt(ind-
1)*dx(1)+J_hs(ind)/2*alpha_hs(1)*dx(1)+K_hs(1)/dx12(1)/2*dx(1)+h1/2*dx
(1);
         c1=-K_hs(1)/dx12(1)/2*dx(1);
         bN=C(end)*dx(end)/2./dt(ind-1)*dx(end)-
J_hs(ind)/2*alpha_hs(end)*dx(end)+K_hs(end)/dx12(end)/2*dx(end)+hN/2*d
x(end);
         aN=-K_hs(end)/dx12(end)/2*dx(end);
     case 3  % Neumann w/ heat capacitance
         h1=bc(4); hN=bc(5); C1=bc(6); CN=bc(7);
         b1=(C(1)*dx(1)/2+C1)./dt(ind-
1)*dx(1)+J_hs(ind)/2*alpha_hs(1)*dx(1)+K_hs(1)/dx12(1)/2*dx(1)+h1/2*dx
(1);
         c1=-K_hs(1)/dx12(1)/2*dx(1);
         bN=(C(end)*dx(end)/2+CN)./dt(ind-1)*dx(end)-
J_hs(ind)/2*alpha_hs(end)*dx(end)+K_hs(end)/2/dx12(end)*dx(end)+hN/2*d
x(end);
         aN=-K_hs(end)/dx12(end)/2*dx(end);
    end
    X(1,1)=b1; X(1,2)=c1; X(end, end-1)=aN; X(end, end)=bN;


    % build the d-vector at interior points
    d=zeros(nx,1);
```

```
    d(2:end-1)=-J_hs(ind)^2*(dx12.^2).*rho(2:end-1)/2-J_hs(ind-
1)^2*(dx12.^2).*rho(2:end-1)/2-...
        T(3:end,ind-1).*K_hs(2:end)./dx2.*dx12/2-...
        T(1:end-2,ind-1).*K_hs(1:end-1)./dx1.*dx12/2+T(2:end-1,ind-
1).*((K_hs(1:end-1)./dx1+...
        K_hs(2:end)./dx2).*dx12/2+J_hs(ind-1)*(alpha_hs(2:end)-
alpha_hs(1:end-1)).*dx12/2-...
        C(2:end-1)./dt(ind-1).*(dx12.^2));

    % now set the boundary conditions in the d vector...
    switch bc(1)
        case 1  % simple Dirichlet
            d1=T1*dx(1); dN=TN*dx(end);
            %  the next two lines add a little heating for the last
dx's
            d(2)=d(2)-(J_hs(ind)^2+J_hs(ind-
1)^2)/2*dx(1)^2*rho(end)/2;
            d(end-1)=d(end-1)-(J_hs(ind)^2+J_hs(ind-
1)^2)/2*dx(end)^2*rho(end)/2;
        case 2  % simple Neumann
            d1=T(1,ind-1)*C(1)*dx(1)/2/dt(ind-1)*dx(1)+h1*T1*dx(1)-
h1/2*T(1,ind-1)*dx(1)+...
                K_hs(1)/dx12(1)/2*dx(1)*(T(2,ind-1)-T(1,ind-1))-...
                J_hs(ind-1)/2*alpha_hs(1)*T(1,ind-1)*dx(1)+...
                (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx(1)^2*rho(1)/2;
            dN=T(end,ind-1)*C(end)*dx(end)/2/dt(ind-
1)*dx(end)+hN*TN*dx(end)-hN/2*T(end,ind-1)*dx(end)+...
                K_hs(end)/dx12(end)/2*dx(end)*(T(end-1,ind-1)-
T(end,ind-1))+...
                J_hs(ind-1)/2*alpha_hs(end)*T(end,ind-1)*dx(end)+...
                (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx(end)^2*rho(end)/2;
        case 3  % Neumann w/ heat capacitance
            d1=T(1,ind-1)*(C1+C(1)*dx(1)/2)/dt(ind-
1)*dx(1)+h1*T1*dx(1)-h1/2*T(1,ind-1)*dx(1)+...
                K_hs(1)/dx12(1)/2*dx(1)*(T(2,ind-1)-T(1,ind-1))-...
                J_hs(ind-1)/2*alpha_hs(1)*T(1,ind-1)*dx(1)+...
                (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx(1)^2*rho(1)/2;
            dN=T(end,ind-1)*(CN+C(end)*dx(end)/2)/dt(ind-
1)*dx(end)+hN*TN*dx(end)-hN/2*T(end,ind-1)*dx(end)+...
                K_hs(end)/dx12(end)/2*dx(end)*(T(end-1,ind-1)-
T(end,ind-1))+...
                J_hs(ind-1)/2*alpha_hs(end)*T(end,ind-1)*dx(end)+...
                (J_hs(ind)^2+J_hs(ind-1)^2)/2*dx(end)^2*rho(end)/2;
    end
    d(1)=d1; d(end)=dN;
 %   keyboard;
    [L,U]=lu(X);  % gets upper and lower triangular factorization
    T(:,ind)=U\(L\d);  % calculate current temperature from previous
one
%    if ind==3
%        keyboard
%    end
```

```
end

% strip off -dt time point used to set up initial conditions
t=t(2:end-1);
T=T(:,2:end);
```

Bibliography

1. A. L. Rockwood, Physical Review A (General Physics) **30** (5), 2843 (1984).
2. Mark Lundstrom, *Fundamentals Of Carrier Transport*, 2ND ED. ed. (CAMBRIDGE UNIV PRESS, Cambridge, 2000).
3. R. Stratton, British Journal of Applied Physics **8** (8), 315 (1957).
4. J. E. Parrott, IEEE Transactions on Electron Devices **43** (5), 809 (1996); K. P. Pipe, R. J. Ram, and A. Shakouri, Physical Review B (Condensed Matter and Materials Physics) **66** (12), 125316 (2002).
5. G. L. Bennett and E. A. Skrabek, in *15th International Conference on Thermoelectrics* (1996), pp. 357; Gary Bennett, in *CRC Handbook of Thermoelectrics*, edited by D. M. Rowe (CRC Press, Boca Raton, FL, 1995), pp. 515.
6. M. Kishi, H. Nemoto, T. Hamao et al., in *18th International Conference on Thermoelectrics* (1999), pp. 301.
7. H. Bottner, in *Proceedings of the 21st International Conference on Thermoelectrics* (2003), pp. 551; J-F. Li, S. Tanaka, T. Umeki et al., Sensors and actuators A **108**, 97 (2003); M. Strasser, R. Aigner, C. Lauterbach et al., Sensors and actuators A **114**, 362 (2004); H. Glosch, M. Ashauer, U. Pfeiffer et al., Sensors and actuators A **74**, 246 (1999).
8. T. Kyono, R. O. Suzuki, and K. Ono, IEEE Transactions on Energy Conversion **18** (2), 330 (2003); A. Tsuyoshi, S. Kagawa, M. Sakamoto et al., in *16th International Conference on Thermoelectrics* (1997), pp. 555.
9. J. W. Stevens, Energy Conversion and Management **42**, 709 (2001).
10. J. G. Haidar and J. I. Ghojel, in *20th International Conference on Thermoelectrics* (2001), pp. 413; D. T. Morelli, in *15th International Conference on Thermoelectrics* (1996), pp. 383.
11. A. C. Alkidas, P. A. Battiston, and D. J. Kapparos, in *Heat Exchangers and Their Simulation, Thermal Management, and Fundamental Advances in Thermal and Fluid Sciences* (Society of Automotive Engineers (SAE), 2004), Vol. SP-1818, pp. 35.
12. Kevin Hoag, *Vehicular Engine Design.* (Springer-Verlag and SAE, Wien, Austria, 2006).
13. G. Min and D. M. Rowe, Energy Conversion and Management **43**, 221 (2002).
14. F. J. Weinberg, D. M. Rowe, and G. Min, J. App. Phys. D. **35**, L61 (2002).
15. J. Vican, B. F. Gajdeczko, F. L. Dryer et al., Proceedings of the Combustion Institute **29**, 909 (2002).
16. R. H. Richman and J. Stringer, in *DARPA/ONR/DOE High Efficiency Thermoelectric Workshop* (2002).
17. Jincan Chen, Zijun Yan, Guoxing Lin et al., Energy Conversion and Management **42** (2), 173 (2001).
18. H. Yugami, H. Sasa, and M. Yamaguchi, Semiconductor Science and Technology **18** (5), 239 (2003).
19. Kuei Fang Hsu, Sim Loo, Fu Guo et al., Science **303** (5659), 818 (2004).
20. H. Goldsmid, in *22nd International Conference on Thermoelectrics* (2003), pp. 433; G. Min and D. M. Rowe, App. Phys. Lett, **77** (6), 860 (2000).
21. T. C. Harman, P. J. Taylor, M. P. Walsh et al., Science **297**, 2229 (2002).
22. R. Venkatasubramanian, E. Silvola, T. Colpitts et al., Nature **413** (6856), 597 (2001).
23. L. D. Hicks and M. S. Dresselhaus, Physical Review B **47**, 12727 (1993).
24. A. Majumdar, Science **303**, 777 (2004).

[25] James Christofferson and Ali Shakouri, Review of Scientific Instruments **76** (2), 024903 (2005).

[26] Dietrich Lueerssen, Janice A. Hudgings, Peter M. Mayer et al., (Institute of Electrical and Electronics Engineers Inc., Piscataway, NJ 08855-1331, United States, San Jose, CA, United States, 2005), pp. 253; G. Tessier, M. L. Polignano, S. Pavageau et al., Journal of Physics D: Applied Physics **39** (19), 4159 (2006).

[27] G. Min and D. M. Rowe, IEE Proc. Sci. Meas. Technol. **143** (6) (1996).

[28] G. Min and D. M. Rowe, in *CRC Handbook of Thermoelectrics*, edited by David M. Rowe (CRC Press, Boca Raton, 1995), pp. 479.

[29] J. M. Gordon and M. Huleihil, J. App. Phys. **72** (3), 829 (1992).

[30] C. Van den Broeck, Physical Review Letters **95** (19), 190602 (2005).

[31] F. L. Curzon and B. Ahlborn, Am. J. Phys. **43**, 22 (1975).

[32] Peter Mayer and Rajeev Ram, Microscale Thermophysical Engineering **10** (2) (2006).

[33] J. P. Fleurial, G. J. Snyder, J. A. Herman et al., (IEEE, Baltimore, MD, USA, 1999), pp. 294.

[34] G. Min and D. M. Rowe, in *21st International Conference on Thermoelectrics* (Long Beach, CA, 2002), pp. 365.

[35] J.-P. Fleurial, G. J. Snyder, J. Patel et al., in *20th International Conference on Thermoelectrics* (2001), pp. 24.

[36] K. D. Hagen, *Heat Transfer with Applications*. (Prentice-Hall, 1999).

[37] X. Fan, G. Zeng, C. LaBounty et al., Applied Physics Letters **78**, 1580 (2001).

[38] D. Allen, R. Haugeto, M. Kajor et al., in *21st International Conference on Thermoelectrics* (2002), pp. 424.

[39] C. B. Vining, in *CRC Handbook of Thermoelectrics*, edited by David M. Rowe (CRC Press, Boca Raton, 1995), pp. 329.

[40] H. B. Callen, Physical Review **73**, 1358 (1948).

[41] G. D. Mahan, 1999; G. D. Mahan, J. O. Sofo, and M. Bartkowiak, Journal of Applied Physics **83** (9), 4683 (1998); Zeng Taofang and Chen Gang, Journal of Applied Physics **92** (6), 3152 (2002); C. B. Vining and G. D. Mahan, Journal of Applied Physics **86** (12), 6852 (1999).

[42] M. D. Ulrich, P. A. Barnes, and C. B. Vining, Journal of Applied Physics **90** (3), 1625 (2001).

[43] L. Helmers, E. Muller, J. Schilz et al., Materials Science and Engineering B **56**, 60 (1998); A. E. Kaliazin, V. L. Kuznetsov, and D. M. Rowe, in *20th International Conference on Thermoelectrics* (2001), pp. 286.

[44] J. L. Cui and X. Qian, (Trans Tech Publications Ltd, Beijing, China, 2003), Vol. 423-425, pp. 367; G. A. Slack and M. A. Hussain, Journal of Applied Physics **70** (5), 2694 (1991); Tongjun Zhang, Jiangyin Peng, Junyou Yang et al., (Beijing, 2001), pp. 303.

[45] G. Jeffrey Snyder, Applied Physics Letters **84** (13), 2436 (2004); G. Jeffrey Snyder and Tristan S. Ursell, Physical Review Letters **91** (14), 148301 (2003).

[46] D. C. Agrawal and V. J. Menon, J. Appl. Phys. D **30**, 357 (1997).

[47] G. Jeffrey Snyder and Tristan Ursell, Physical Review Letters **91** (14), 148301 (2003).

[48] F. P. Incropera, *Liquid cooling of electronic devices by single-phase convection*. (John Wiley and Sons, 1999).

[49] O. Yamashita and N. Sadatomi, Journal of Applied Physics **88**, 245 (2000).

[50] A. Jacquot, W. L. Liu, G. Chen et al., in *21st International Conference on Thermoelectrics* (2002), pp. 118.

[51] Cronin B. Vining, Journal of Applied Physics **69** (1), 331 (1991).

[52] D. J. Eaglesham, Journal of Applied Physics **77** (8), 3597 (1995); M. A. Hermanlk and H. Sitter, Microelectronics Journal **27** (4-5), 257 (1996); J. W. Matthews and A. E. Blakeslee, Journal of Crystal Growth **27** (1), 118 (1974).

[53] Lee, Song, Van Au et al., in *ASME/JSME Thermal Engineering Conference* (1995), Vol. 4, pp. 199.

[54] D. Lorenzen, J. Bonhaus, W. R. Fahrner et al., IEEE Transactions on Industrial Electronics **48** (2), 286 (2001).

[55] Denis Maillet, *Thermal quadrupoles: solving the heat equation through integral transforms.* (Wiley, Chichester; New York, 2000).

[56] C. Garrod and J. Hurley, Physical Review A (General Physics) **27** (3), 1487 (1983).

[57] D. G. Cahill, Review of Scientific Instruments **61** (2), 802 (1990).

[58] Raymond Jay Emrich, *Fluid dynamics.* (Academic Press, New York, 1981).

[59] S. Song, K. P. Moran, D. P. Rearick et al., San Diego, CA, USA, 1993 (unpublished).

[60] A. Miner and U. Ghoshal, Applied Physics Letters **85**, 506 (2004).

[61] Koo Jae-Mo, Im Sungjun, Jiang Linan et al., Transactions of the ASME. Journal of Heat Transfer **127** (1), 49 (2005).

[62] D. B. Tuckerman and R. F. W. Pease, IEEE Electron Device Letters **EDL-2** (5), 126 (1981).

[63] Ray Beach, William J. Benett, B. L. Freitas et al., IEEE Journal of Quantum Electronics **28** (4), 966 (1992); K. E. Goodson, K. Kurabayashi, and R. F. W. Pease, (ASME, New York, NY, USA, Portland, OR, USA, 1997), Vol. 305, pp. 187; Sanjay K. Roy and Branko L. Avanic, IEEE Transactions on Components, Packaging, and Manufacturing Technology Part B: Advanced Packaging **19** (2), 444 (1996); J. A. Skidmore, B. L. Freitas, J. Crawford et al., Applied Physics Letters **77** (1), 10 (2000).

[64] R.J. Phillips, *Microchannel Heat Sinks.* (ASME, New York, 1990).

[65] C. Y. Zhao and T. J. Lu, International Journal of Heat and Mass Transfer **45** (24), 4857 (2002).

[66] Lindon C. Thomas, *Heat transfer: professional version*, 2nd ed. (Capstone Pub., Tulsa, Okla., 1999).

[67] A. F. Mills, *Basic Heat and Mass Transfer.* (Irwin, Chicago, 1995); C. G. du Toit, Numerical Heat Transfer: Part B: Fundamentals **41** (3-4), 397 (2002).

[68] D. M. Rowe, *CRC handbook of thermoelectrics.* (CRC Press, Boca Raton, FL, 1995).

[69] D. Langridge, W. Lawrance, and B. Wichert, Solar Energy **56** (2), 151 (1996).

[70] S. G. Kandlikar and H. R. Upadhye, San Jose, CA, USA, 2005 (unpublished).

[71] Qu Weilin and I. Mudawar, International Journal of Heat and Mass Transfer **45** (12), 2549 (2002).

[72] R. J. Phillips, Lincoln Laboratory Journal **1** (1), 31 (1988).

[73] Pega Hrnjak and Xiao Tu, International Journal of Heat and Fluid Flow **28** (1 SPEC ISS), 2 (2007).

[74] S. Del Giudice, C. Nonino, and S. Savino, International Journal of Heat and Fluid Flow **28** (1 SPEC ISS), 15 (2007).

[75] T. C. Harman, J. H. Cahn, and M. J. Logan, Journal of Applied Physics **30** (9), 1351 (1959); G. Gromov, D. Kondratiev, A. Rogov et al., in *Sixth European Workshop on Thermoelectricity of the European Thermoelectric Society* (Freiburg, 2001), pp. 1; Richard J. Buist, in *CRC Handbook of Thermoelectrics*, edited by David M. Rowe (CRC Press, Boca Raton, 1995), pp. 189.

[76] A. W. Penn, Journal of Scientific Instruments **41** (10), 626 (1964).

[77] A. Green and L. E. J. Cowles, Journal of Scientific Instruments **37** (9), 349 (1960).

[78] Rajeev Singh, Bian Zhixi, Zeng Gehong et al., in *Materials Research Symposium* (Boston, 2005).

[79] Hugh H. Woodbury, Lionel M. Levinson, and Robert S. Lewandowski, in *CRC Handbook of Thermoelectrics*, edited by David M. Rowe (CRC Press, Boca Raton, 1995), pp. 181.

[80] E. Muller, J. U. Bruch, and J. Schilz, (Trans Tech Publications, Dresden, Germany, 1999), Vol. 308-311, pp. 754.

[81] Jim J. W. Tzeng, Tom W. Weber, and Dan W. Krassowski, Annual IEEE Semiconductor Thermal Measurement and Management Symposium, 174 (2000).
[82] David J. Cahill, Henry E. Fischer, Tom Klitsner et al., J. Vac. Sci. Technol. A **7** (3), 1259 (1989).
[83] B. Yang, J. L. Liu, K. L. Wang et al., in *20th International Conference on Thermoelectrics* (IEEE, 2001), pp. 344.
[84] T. Borca-Tasciuc, A. R. Kumar, and G. Chen, Review of Scientific Instruments **72** (4), 2139 (2001).
[85] W. L. Liu, T. Borca-Tasciuc, J. L. Liu et al., (IEEE, Beijing, China, 2001), pp. 340.
[86] L. K. Mak, C. M. Rogers, and D. C. Northrop, Journal of Physics E: Scientific Instruments **22** (5), 317 (1989).
[87] B. Yang, W. L. Liu, J. L. Liu et al., Applied Physics Letters **81** (19), 3588 (2002).
[88] M. A. M. Gijs, J. B. Giesbers, S. K. J. Lenczowski et al., App. Phys. Lett, **63** (1), 111 (1993).
[89] Zhang Yan, Zeng Gehang, R. Singh et al., (IEEE, Long Beach, CA, USA, 2002), pp. 329; Zhang Yan, D. Vashaee, R. Singh et al., (Mater. Res. Soc, Boston, MA, USA, 2004), pp. 59.
[90] Zeng Gehong, J. E. Bowers, Zhang Yan et al., Clemson, SC, USA, 2005 (unpublished).
[91] E. Muller, J. U. Bruch, and J. Schilz, (IEEE, Nagoya, Japan, 1998), pp. 441.
[92] John M. Damaschke, IEEE Transactions on Industry Applications **33** (5), 1203 (1997).
[93] E. G. Wolff and D. A. Schneider, International Journal of Heat and Mass Transfer **41** (22), 3469 (1998).
[94] Zeng Taofang and Chen Gang, Transactions of the ASME. Journal of Heat Transfer **123** (2), 340 (2001).
[95] J. W. Sharp, E. H. Volckmann, H. B. Lyon, Jr. et al., (IEEE, Pasadena, CA, USA, 1996), pp. 417.
[96] H. Scherrer and S. Scherrer, in *CRC Handbook of Thermoelectrics*, edited by David M. Rowe (CRC Press, Boca Raton, 1995), pp. 211.
[97] S. Berbenni, V. Favier, and M. Berveiller, International Journal of Plasticity **23** (1), 114 (2007).
[98] N. Keawprak, Z. M. Sun, H. Hashimoto et al., Journal of Alloys and Compounds **397** (1-2), 236 (2005).
[99] Yan Zhang, Ali Shakouri, and Gehong Zeng, Applied Physics Letters **85** (14), 2977 (2004).
[100] S. Chakraborty, C. A. Kleint, A. Heinrich et al., Applied Physics Letters **83** (20), 4184 (2003).
[101] G. Chen, B. Yang, W. L. Liu et al., (IEEE, Beijing, China, 2001), pp. 30.
[102] Zeng Gehong, J. E. Bowers, Zhang Yan et al., (IEEE, Clemson, SC, USA, 2005), pp. 500.
[103] J. M. Zide, D. O. Klenov, S. Stemmer et al., Applied Physics Letters **87** (11), 112102 (2005).
[104] A. Dorn, M. Peter, S. Kicin et al., Applied Physics Letters **82** (16), 2631 (2003).
[105] Woochul Kim, Joshua Zide, Arthur Gossard et al., Physical Review Letters (96), 045901 (2006).
[106] D. L. Rode, Physical Review B (Solid State) **3** (10), 3287 (1971).
[107] S. A. Aliev, A. Ya Nashel'skii, and S. S. Shalyt, Fizika Tverdogo Tela **7** (5), 1590 (1965).
[108] P. Mayer and R. J. Ram, (IEEE, Clemson, SC, USA, 2005), pp. 280.
[109] R. Mansfield and W. Williams, Proceedings of the Physical Society **72**, 733 (1958).
[110] Peter Y. Yu and Manuel Cardona, *Fundamentals of semiconductors: physics and materials properties*, 3rd, rev. and enlarged ed. (Springer, Berlin New York, 2001); E. Matatagul and M. Cardona, Solid State Communications **6** (5), 313 (1968).

[111] E. Matatagui, A. G. Thompson, and M. Cardona, Physical Review **176** (3), 950 (1968).

[112] Shunji Ozaki and Sadao Adachi, Journal of Applied Physics **78** (5), 3380 (1995).

[113] P. B. Allen and M. Cardona, Physical Review B (Condensed Matter) **23** (4), 1495 (1981); P. B. Allen and M. Cardona, Physical Review B (Condensed Matter) **27** (8), 4760 (1983); P. B. Allen and V. Heine, Journal of Physics C (Solid State Physics) **9** (12), 2305 (1976); Y. P. Varshni, Physica **34** (1), 149 (1967); P. Lautenschlager, M. Garriga, and M. Cardona, Physical Review B (Condensed Matter) **36** (9), 4813 (1987); P. Lautenschlager, M. Garriga, L. Vina et al., Physical Review B (Condensed Matter) **36** (9), 4821 (1987).

[114] B. Batz, Solid State Communications **5** (12), 985 (1967).

[115] L. Vina, S. Logothetidis, and M. Cardona, Physical Review B (Condensed Matter) **30** (4), 1979 (1984).

[116] B. O. Seraphin and N. Bottka, Physical Review **145** (2), 628 (1966).

[117] Shunji Ozaki, Sadao Adachi, Masahiro Sato et al., Journal of Applied Physics **79** (1), 439 (1996).

[118] Takayuki Miyazaki and Sadao Adachi, Journal of Applied Physics **77** (4), 1741 (1995).

[119] Sadao Adachi, Shunji Ozaki, Masahiro Sato et al., Japanese Journal of Applied Physics, Part 1: Regular Papers & Short Notes & Review Papers **35** (2A), 537 (1996); Ching-Hwa Ho, Horng-Wen Lee, and Zau-Hwang Cheng, Review of Scientific Instruments **75** (4), 1098 (2004).

[120] B. M. Clemens, G. L. Eesley, and C. A. Paddock, Physical Review B (Condensed Matter) **37** (3), 1085 (1988); G. L. Eesley, Physical Review B (Condensed Matter) **33** (4), 2144 (1986); A. Miklos and A. Lorincz, Journal of Applied Physics **63** (7), 2391 (1988); N. Taketoshi, T. Baba, and A. Ono, Measurement Science & Technology **12** (12), 2064 (2001); Andrew N. Smith and Pamela M. Norris, Applied Physics Letters **78** (9), 1240 (2001).

[121] C. A. Paddock and G. L. Eesley, Journal of Applied Physics **60** (1), 285 (1986); M. G. Burzo, P. L. Komarov, and P. E. Raad, IEEE Transactions on Components and Packaging Technologies **28** (1), 39 (2005); Mihai G. Burzo, Pavel L. Komarov, and Peter E. Raad, IEEE Transactions on Components and Packaging Technologies **26** (1), 80 (2003); Pavel L. Komarov, Peter E. Raad, and M. G. Burzo, Journal of Heat Transfer **124** (6), 1009 (2002); David G. Cahill, Review of Scientific Instruments **75** (12), 5119 (2004).

[122] Bincheng Li, L. Pottier, J. P. Roger et al., Review of Scientific Instruments **71** (5), 2154 (2000); Li Bincheng, J. P. Roger, L. Pottier et al., Journal of Applied Physics **86** (9), 5314 (1999); G. Langer, J. Hartmann, and M. Reichling, Review of Scientific Instruments **68** (3), 1510 (1997); J. Hartmann, M. Costello, and M. Reichling, Physical Review Letters **80** (1), 117 (1998); M. Reichling, T. Klotzbucher, and J. Hartmann, Applied Physics Letters **73** (6), 756 (1998); J. Hartmann, P. Voigt, and M. Reichling, Journal of Applied Physics **81** (7), 2966 (1997); A. A. Maznev, J. Hartmann, and M. Reichling, Journal of Applied Physics **78** (9), 5266 (1995); H. Gronbeck and M. Reichling, Journal of Applied Physics **78** (11), 6408 (1995); M. Reichling and H. Gronbeck, Journal of Applied Physics **75** (4), 1914 (1994).

[123] P. Voigt, J. Hartmann, and M. Reichling, J. App. Phys. **80** (4), 2013 (1996).

[124] C. Filloy, G. Tessier, S. Hole et al., Sensor Review **23** (1), 35 (2003).

[125] G. Tessier, S. Hole, and D. Fournier, Applied Physics Letters **78** (16), 2267 (2001).

[126] Y. Li, Y. Lu, H. Shen et al., Applied Physics Letters **70** (18), 2458 (1997).

[127] G. Tessier, G. Jerosolimski, S. Hole et al., Review of Scientific Instruments **74** (1 II), 495 (2003).

[128] Gilles Tessier, Stephane Hole, and Daniele Fournier, Optics Letters **28** (11), 875 (2003).

[129] B. R. Bennett, R. A. Soref, and J. A. Del Alamo, IEEE Journal of Quantum Electronics **26** (1), 113 (1990).
[130] Tatiana Bienville, Laurent Belliard, Pierre Siry et al., (Academic Press, London, United Kingdom, Reims, France, 2004), Vol. 35, pp. 363.
[131] R. Cherrak, V. Loriette, B. Forget et al., (The International Society for Optical Engineering, Boulder, CO, United States, 1997), Vol. 3244, pp. 570; J. Opsal, A. Rosencwaig, and D. L. Willenborg, Applied Optics **22** (20), 3169 (1983); G. Busse, D. Wu, and W. Karpen, Journal of Applied Physics **71** (8), 3962 (1992).
[132] S. Dilhaire, Y. Ezzahri, S. Grauby et al., in *22nd International Conference on Thermoelectrics* (2003), pp. 519.
[133] Christopher J. Stolz, Diane J. Chinn, Robert D. Huber et al., (International Society for Optical Engineering, Bellingham, WA 98227-0010, United States, Boulder, CO, United States, 2004), Vol. 5273, pp. 373.
[134] A. Dubois, A. C. Boccara, and M. Lebec, Optics Letters **24** (5), 309 (1999); T. Spirig, P. Seitz, O. Vietze et al., IEEE Journal of Quantum Electronics **31** (9), 1705 (1995).
[135] James Christofferson and Ali Shakouri, (Institute of Electrical and Electronics Engineers Inc., San Jose, CA., United States, 2004), Vol. 20, pp. 87.
[136] James Christofferson and Ali Shakouri, (Institute of Electrical and Electronics Engineers Inc., San Jose, CA, United States, 2003), pp. 271; J. Christofferson, D. Vashaee, A. Shakouri et al., (IEEE, Monterey, CA, USA, 2000), pp. 49; J. Christofferson, D. Vashaee, A. Shakouri et al., (SPIE-Int. Soc. Opt. Eng, San Jose, CA, USA, 2001), Vol. 4275, pp. 119.
[137] J. Christofferson, D. Vashaee, A. Shakouri et al., (IEEE, San Jose, CA, USA, 2001), pp. 58.
[138] S. Grauby, B. C. Forget, S. Hole et al., Review of Scientific Instruments **70** (9), 3603 (1999).
[139] M. Von Arx, O. Paul, and H. Baltes, IEEE Transactions on Semiconductor Manufacturing **10** (2), 201 (1997).
[140] Stefan Dilhaire, Stephane Grauby, Sebastien Jorez et al., Applied Optics **41** (24), 4996 (2002).
[141] T. Phan, S. Dilhaire, V. Quintard et al., Measurement Science & Technology **8** (3), 303 (1997); S. Grauby, S. Dilhaire, S. Jorez et al., Review of Scientific Instruments **74** (1 II), 645 (2003).
[142] S. Dilhaire, S. Grauby, and W. Claeys, IEEE Electron Device Letters **26** (7), 461 (2005).
[143] Dietrich Luerssen, Janice A. Hudgings, Peter M. Mayer et al., San Jose, CA, United States, 2005 (unpublished).
[144] O. Breitenstein and M. Langenkamp, *Lock-in thermography: basics and use for functional diagnostics of electronic components.* (Springer, New York, 2003).
[145] K. J. Ito, Y. Uehara, S. Ushioda et al., Review of Scientific Instruments **71** (2 pt 1), 420 (2000).
[146] M. Fujinami, K. Toya, and T. Sawada, Review of Scientific Instruments **74** (1), 621 (2003).
[147] S. Grauby, G. Tessier, V. Rachet et al., Analytical Sciences **17**, 67 (2001).
[148] L. Gammaitoni, P. Hanggi, P. Jung et al., Reviews of Modern Physics **70** (1), 223 (1998).
[149] S. B. Ippolito, S. A. Thorne, M. G. Eraslan et al., Applied Physics Letters **84** (22), 4529 (2004).
[150] David Cahill, Kenneth Goodson, and Arunava Majumdar, Journal of Heat Transfer **124**, 223 (2002).
[151] L. Shi, O. Kwon, A. C. Miner et al., Journal of Microelectromechanical Systems **10** (3), 370 (2001); Shi Li, S. Plyasunov, A. Bachtold et al., Applied Physics Letters **77** (26),

4295 (2000); A. Majumdar, K. Luo, Z. Shi et al., Experimental Heat Transfer **9** (2), 83 (1996).

152 J. C. Mullikin, L. J. Van Vliet, H. Netten et al., in *Image acquisition and scientific imaging systems*, edited by H. C. Titus and A. Waks (1994), Vol. 2173, pp. 73; Y. Reibel, M. Jung, M. Bouhifd et al., The European Physical Journal of Applied Physics **21**, 75 (2003).

153 R. J Wannamaker, S. P. Lipshitz, J. Vanderkooy et al., IEEE Transactions on Signal Processing **48** (2), 499 (2000).

154 Haykin, *Digital Communications*. (Wiley, NY, 1988); M. D. Kulkarni and A. B. Kostinski, IEEE Transactions on Geoscience and Remote Sensing **33** (3), 799 (1995).

155 C. Benard, Y. Body, M. Delisee et al., International Journal of Heat and Mass Transfer **33** (5), 785 (1990); E. Schreck, B. Hiller, and G. P. Singh, Review of Scientific Instruments **64** (1), 218 (1993).

156 Clyde F. Coombs, *Electronic instrument handbook*, 3rd ed. (McGraw-Hill, New York, 2000).

157 Peter Mayer and Rajeev Ram, in *Mater. Res. Soc. Symp. Proc.* (Boston, Fall 2005), Vol. 886.

158 S. Dilhaire, S. Grauby, and W. Claeys, Applied Physics Letters **84** (5), 822 (2004).

159 K. Taniguchi, A. Moritani, C. Hamaguchi et al., Surface Science **37**, 212 (1973).

160 D. L. Greenaway and G. Harbeke, Journal of the Physics and Chemistry of Solids **26**, 1585 (1965); S. Katsuki, Journal of the Physical Society of Japan **26** (1), 58 (1969).

161 Dietrich Lueerssen, Peter Mayer, Janice Hudgings et al., presented at the Semitherm, San Jose, 2005 (unpublished).

162 Marlow Industries.

163 Adam Downey, Edward Timm, Pierre Poudeu et al., in *Materials Research Symposium (MRS)* (Boston, 2005).

164 T. C. Harman, M. P. Walsh, B. E. Laforge et al., Charlotte, NC, USA, 2005 (unpublished).

165 R. Lake and S. Datta, Physical Review B (Condensed Matter) **46** (8), 4757 (1992).

166 D. Vashaee and A. Shakouri, Journal of Applied Physics **95** (3), 1233 (2004); D. Vashaee and A. Shakouri, Physical Review Letters **92** (10), 106103 (2004).

167 A. Wacker, A. P. Jauho, S. Rott et al., Physica B: Condensed Matter **272** (1), 157 (1999).

168 M. Zebarjadi, A. Shakouri, and K. Esfarjani, Physical Review B (Condensed Matter and Materials Physics) **74** (19), 195331 (2006).

169 Annie Wang, Ioannis Kymissis, Vladimir Bulovic et al., Applied Physics Letters **89** (11), 112109 (2006); Ioannis Kymissis, Akintunde Ibitayo Akinwande, and Vladimir Bulovic, IEEE/OSA Journal of Display Technology **1** (2), 289 (2005); Annie Wang, Ioannis Kymissis, Vladimir Bulovic et al., San Francisco, CA, United States, 2004 (unpublished).

170 K. Ryu, I. Kymissis, V. Bulovic et al., IEEE Electron Device Letters **26** (10), 716 (2005).

171 Roxann R. Blanchard and J. A. del Alamo, IEEE Transactions on Electron Devices **53** (6), 1289 (2006).

172 A. A. Villanueva, J. A. Del Alamo, T. Hisaka et al., Washington, DC, United States, 2003 (unpublished).

173 J. A. Del Alamo and A. A. Villanueva, San Francisco, CA, United States, 2004 (unpublished).