# Thermoelectric Properties of Bismuth and Silicon Nanowires

Thesis by

Akram Issam Boukai

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2008

(Defended January 16, 2008)

© 2008

Akram Issam Boukai

*Dedicated in memory of my grandmother Neda, for her hard work,*

*generosity, and determination which represents the best of Arab peoples.*

## Acknowledgements

I would like to give many thanks to my advisor Jim Heath whom I have been working with for almost nine years. Many good things can be said about Jim, such as his intensity and dedication to science which rubs off on his students. I would like to point out my favorite quality of his which can be summarized through a quote by Mark Twain: "…the really great make you feel that you, too, can become great". I am lucky to have met Jim. He is truly a great person.

I would also like to thank the people on my thesis committee, Bill Goddard, Nate Lewis, and Harry Atwater. Bill is an amazing individual whose sincerity and open-mindedness allowed for a great collaboration on the silicon thermoelectric project. Bill's drive to understand our experimental results was infectious. Nate and Harry have been extremely valuable to my growth as a scientist, especially the grilling they gave me at my $5^{th}$ year proposition exam. Thank you for caring!

The people that I met in the Heath group have all been unique and my time with them has been rewarding. I would especially like to thank Dr. Yi Luo, Dr. Dunwei Wang, Dr. Nick Melosh, and Dr. Mike Diehl. Their guidance over the years helped me to grow into the scientist I am today and for that I am truly grateful. Special thanks must go to Dr. Jamil Tahir-Kheli whose brilliance is unmatched. His contribution to the theory section of the silicon thermoelectric paper was extraordinary. Additional thanks go to my friend Dr. Jamie Link whose camaraderie and wit are sorely missed. I am going to miss Caltech because people like them make Caltech special.

I would like to thank Ke Xu, and Yuri Bunimovich and Jen-Kan Yu for their help on the bismuth and silicon thermoelectric projects respectively. Also, my office mates

and friends, John Nagarah, Jonathan Green, and Habib Ahmad deserve mention for making my time in the office enjoyable. Thanks go to Douglas Tham who is going to be a super star and to Dr. Michael McAlpine whose company was always entertaining. I would also like to thank Bonnie Sheriff and Dr. Erica DeIonno for being wonderful lab mates.

Last but not least, I would like to thank my loved ones, especially my mom, dad, my sister Douda, and fiancé Chioma. My mom and dad have taught me the value of hard work and lead by example for which I am forever indebted. The sacrifices they made for my sister and I will never be forgotten. My sister is truly an inspiration and I am very proud of her accomplishments. My fiancé Chioma is the woman of my dreams and I am fortunate to have found her. I thank them all for their endless support.

## Abstract

Thermoelectric materials convert temperature differences into electricity and vice versa. Such materials utilize the Seebeck effect for power generation and the Peltier effect for refrigeration. In the Seebeck effect, a temperature gradient across a material causes the diffusion of charged carriers across that gradient, thus creating a voltage difference between the hot and cold ends of the material. Conversely, the Peltier effect explains the fact that when current flows through a material a temperature gradient arises because the charged carriers exchange thermal energy at the contacts. Thermoelectrics perform these functions without moving parts and they do not pollute. This makes them highly reliable and more importantly attractive as renewable energy sources, especially at a time when global warming is a growing concern. However, thermoelectrics find only limited use because of their poor efficiency.

The efficiency of a thermoelectric material is determined by the dimensionless figure of merit, $ZT = \dfrac{S^2 \sigma}{\kappa} T$, where $S$ is the thermoelectric power, defined as the thermoelectric voltage, $V$, produced per degree temperature difference $S = \dfrac{\mathrm{d}V}{\mathrm{d}T}$, $\sigma$ is the electrical conductivity, $\kappa$ is the thermal conductivity, and $T$ is the temperature. To maximize $ZT$, $S$ must be large so that a small temperature difference can create a large voltage, $\sigma$ must be large in order to minimize joule heating losses, and $\kappa$ must be small to reduce heat leakage and maintain a temperature difference. Maximizing $ZT$ is challenging because optimizing one physical parameter often adversely affects another. The best commercially available thermoelectric devices are alloys of $Bi_2Te_3$ and have a $ZT$ of 1 which corresponds to a carnot efficiency of ~10%. My research has focused on

achieving efficient thermoelectric performance from the single component systems of bismuth and silicon nanowires.

Bismuth nanowires are predicted to undergo a semi-metal to semiconductor transition below a size of 50 nm which should increase the thermopower and thus $ZT$. Limited experimental evidence by other groups has been acquired to support this claim. Through electric field gating measurements and by tuning the nanowire size, we have shown that no such transition occurs. Instead, surface states dominate the electric transport at a size smaller than 50 nm and bismuth remains a semimetal.

Bulk silicon is a poor thermoelectric due to its large thermal conductivity. However, silicon nanowires may have a dramatically reduced thermal conductivity. By varying the nanowire size and impurity doping levels, $ZT$ values representing an approximately 100-fold improvement over bulk silicon are achieved over a broad temperature range, including a $ZT \sim 1$ at 200K. Independent measurements of $S$, $\sigma$, and $\kappa$, combined with theory, indicate that the improved efficiency originates from phonon effects. The thermal conductivity is reduced and the thermopower is enhanced. These results are expected to apply to other classes of semiconductor nanomaterials.

# Table of Contents

# List of Figures

**Chapter 1**

**Chapter 2**

**Chapter 3**

**Chapter 4**

**Chapter 5**

# Chapter 1

# Thermoelectrics: History and Physics

## 1.1    Introduction

In 1821 Thomas Johann Seebeck discovered that a magnetic compass would deflect in the presence of a material with ends at different temperatures[1].  He attributed the effect to magnetism.   Seebeck's interpretation was wrong.   The correct reason was that the temperature difference created a current in the material that deflected the magnet.  During seminars, he would vigorously defend his position in light of mounting evidence that suggested the effect was due to an electric current alone that was generated by the temperature difference.  Seebeck's discovery, however, would later be named after him. In science, sometimes the interpretation of a discovery is not as important as the discovery itself.

Seebeck's discovery that a temperature gradient across a material creates a voltage was later followed by Peltier's discovery in 1834 of the reverse effect.  Peltier showed that a current through a material can freeze water at one end and melt ice at the other[2].   Peltier and Seebeck had discovered two physical effects that had the same underlying physics, and that could do useful work.  The Seebeck effect can be used to generate electric power while the Peltier effect can be used to provide refrigeration. However, it wasn't until William Thompson's (Lord Kelvin) discovery of a third effect that the physics of thermoelectricity began to be understood.   The Thompson effect

described a material evolving or absorbing heat under the presence of both a current and temperature gradient.  Through an educated guess, Thompson correctly predicted that Seebeck's and Peltier's discoveries were actually different manifestations of the same underlying physics.  In 1931, Onsager rigorously proved Thompson's educated guess and set the stage for a flurry of experimental studies that were focused on increasing the thermoelectric efficiency.

## 1.2    Thermoelectric Efficiency

Thermoelectric devices offer the promise of power generation and refrigeration with no moving parts and nearly maintenance free operation[3-5].  A temperature difference across any non-superconducting material will create a voltage that can be used to provide power.  A current through that same material can be used to cool one end for refrigeration purposes.  Thermoelectrics are essentially particle heat engines where the working fluid is composed of either electrons or holes.  The outstanding challenge in the study of thermoelectrics is to find materials that can provide power and refrigeration at a high thermodynamic efficiency.  Presently, thermoelectrics are thirty to twenty percent less efficient than present day fossil-fuel burning power plants and compressor-based refrigerators.  The thermoelectric principles are well understood for bulk materials and the most efficient ones are based on alloys of $Bi_2Te_3$ with efficiencies[6] hovering around 10%.

The efficiency of a thermoelectric material is determined by the dimensionless figure of merit, $ZT = \dfrac{S^2 \sigma}{\kappa} T$, where $S$ is the thermoelectric power, defined as the

thermoelectric voltage ($V$) produced per degree temperature difference $S = \dfrac{\Delta V}{\Delta T}$, $\sigma$ is the electrical conductivity, $\kappa$ is the total thermal conductivity consisting of both phonon and electronic contributions, and $T$ is the temperature. The efficiency is related to $ZT$ by the following equation[7] (see Appendix 1.1 for a full derivation):

$$\eta = \frac{\Delta T}{T_{Hot}} \frac{\sqrt{1+ZT} - 1}{\sqrt{1+ZT} + \dfrac{T_{Cold}}{T_{Hot}}} \quad \ldots\ldots\ldots\text{equation 1}$$
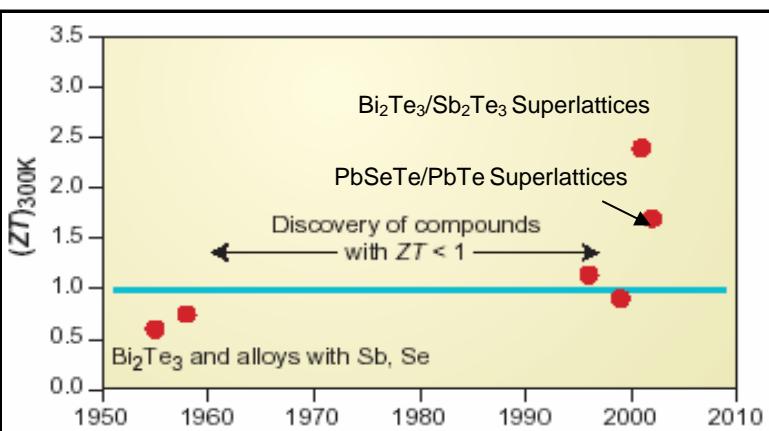


**Figure 1-1. Progression of the Figure of Merit, *ZT*, During the Last Fifty Years.** *ZT* remained nearly stagnant at around 1 from the 1950's to the '90s until experiments on nanostructured materials pushed the value to nearly 2.5 at the start of the millennium. Reprinted from Ref. 8.

where $T_{Hot}$ is the temperature of the hot junction and $T_{cold}$ is the temperature of the cold junction. One immediately sees from equation 1 that the efficiency tends towards the Carnot limit when $ZT$ approaches infinity. To compete with modern-day compressor based refrigerators, a thermoelectric device must have a $ZT > 3$ (Figure 1-1). There is no theoretical limit to $ZT$. To maximize $ZT$, $S$ must be large so that a small temperature difference can create a large voltage, $\sigma$ must be large in order to minimize joule heating losses, and $\kappa$ must be small to reduce heat leakage and maintain a temperature difference[9]. $S$, $\sigma$, and $\kappa$ are not, however, mutually exclusive. For instance in a bulk material, the Weidmann-Franz law limits the ratio $\sigma/\kappa_e$, where $\kappa_e$ is the electronic contribution to the total thermal

conductivity. Also, a sharply peaked density of states favors large $S$, but the density of states is a smoothly varying function for bulk materials. Nanowires may offer the possibility to circumvent some of the materials limitations found in bulk. First, the thermal conductivity has been shown to decrease when the size of a material is reduced to the nano-scale[10-13]. Second, the thermopower is sensitive to the density of states and phonon-electron scattering, which may be very different in certain materials at the nano-scale[14]. To increase $ZT$, a better understanding of the interplay between the three materials parameters, $S$, $\sigma$, and $\kappa$, must be developed.

## 1.3    Fundamental Principles—The Seebeck Effect for Power Generation

The Seebeck, thermoelectric power, or thermopower, effect describes the fact that a temperature gradient drives the diffusion of charged carriers which in turn creates a voltage[3-5]. The Seebeck coefficient of a material depends on the material's temperature and crystal structure. The terms thermoelectric power and thermopower are misnomers since they measure the voltage or electric field (not the electric power) induced in response to a temperature difference. Fundamentally, the thermopower is a measure of the amount of net entropy transported per carrier[15]. The major contributors to the transfer of entropy in the presence of a temperature gradient originate from diffusion and/or phonon drag effects. Both effects are seen in metals and semiconductors, however, one typically dominates over the other. For example, diffusion dominates the thermopower for most metals whereas phonon drag can be a hundred times larger than diffusion in semiconductors.

## 1.3.1 Diffusion Thermopower

Classically, gas molecules that are heated possess larger average velocities than when cold and thus diffuse more rapidly. In the solid state, a mobile electron gas diffusing to the cold side leaves behind the oppositely charged and immobile nuclei at the hot side which gives rise to a thermoelectric voltage. This separation of charges creates an electric field. At open-circuit equilibrium the buildup of electrons onto the cold side from diffusion eventually ceases at some maximum value because of the oppositely directed drift of charged carriers from the electric force-field. Only an increase in the temperature difference can resume the buildup of more charges on the cold side and thus lead to an increase in the thermoelectric voltage.

A classical explanation of the Seebeck effect only explains the thermoelectric voltage due to electrons and ignores holes. Also, classically, the statistics of the particles is ignored but electrons and holes are fermions and must obey Fermi-Dirac statistics. In reality, the sign of the thermoelectric voltage can take two values, positive or negative, depending on the dominant charge carriers—electrons or holes. By convention, if the dominant charge carriers are electrons then the Seebeck coefficient is negative and vice versa for holes.

## 1.3.2 Diffusion Thermopower in Metals

The analysis of the thermopower in a metal can be simplified if it is assumed that the density of states is that of a 3-D free electron system. Only electrons near the Fermi surface participate in the electronic transport. The electrons in such a material possess $kT$ energy and there are $kT/E_F$ electrons with that energy where k is Boltzmann's constant, $T$

is the temperature and $E_F$ is the Fermi energy. The thermopower is the amount of entropy transported per electronic charge. *kT/T* is the entropy the electrons transfer where *kT* is the amount of heat they possess. Therefore the thermopower of a 3-D free electron system is proportional to $\dfrac{k^2 T}{e E_F}$ where *e* is the electronic charge.

A more accurate quantum-mechanical explanation of the Seebeck effect was given by Sir Neville Mott derived from Boltzmann's transport equation[16]. The Mott equation, valid only for metals, is:

$$S = \frac{\pi^2 k^2 T}{3e} \frac{\partial \ln \sigma(E)}{\partial E}\bigg|_{E_F} \qquad \ldots\ldots\ldots\text{equation 2}$$

where *T* is the temperature, *k* is Boltzmann's constant, e is the electronic charge, and $\sigma$ is the conductivity when the electron energy, *E*, equals to the Fermi energy, $E_F$. In a metal the electrical conductivity is a function of the density of states, scattering rate, and carrier velocity. Therefore equation 1 can be written:

$$S = \frac{\pi^2 k^2 T}{3e} \left[ \frac{\partial \ln N(E)}{\partial E}\bigg|_{E_F} + \frac{\partial \ln \tau(E)}{\partial E}\bigg|_{E_F} + \frac{\partial \ln v^2(E)}{\partial E}\bigg|_{E_F} \right] \qquad \ldots\ldots\text{equation 3}$$

where *N(E)* is the density of states (DOS), $\tau$ is the scattering rate and *v* is the carrier velocity. The Seebeck coefficient is therefore independent of the magnitude of these three terms and depends only on their rate of change. The diffusion thermopower is linearly dependent on *T* and due to the second law of thermodynamics goes to zero near 0K. A thermodynamic treatment of the change in chemical potential upon application of

a temperature gradient that ignores carrier scattering leads to a more intuitive explanation for the Seebeck coefficient.

### 1.3.3 The (Electro-)Chemical Potential

The chemical potential is defined as the amount of energy needed to add one particle into a system while keeping the entropy and temperature constant. The chemical potential is a measure of the energy brought into and out of a system when particles are exchanged, akin to the temperature which measures the energy change of a system when entropy is exchanged[17]. Particles flow or diffuse when there are chemical potential differences. Chemical potential differences are responsible for the operation of most power generators such as batteries, solar cells, and thermoelectrics.

**Batteries**. Batteries provide power by employing anodes and cathodes that have different chemical potentials for electrons. Low lying energy states in the cathode cause electrons to flow from anode to cathode when the two are connected, for example, by a metallic wire. The flow of electrons from anode to cathode would cease if not for the electrolyte which contains chemicals that either give up electrons or accept them, thus replenishing the anode with electrons and removing them from the cathode. When the chemicals are used up, the battery can no longer provide power. The chemical potential difference between anode and cathode, therefore, is the cause of the electrical current flow. If the anode and cathode are the same material, then there would be no net chemical potential difference and the battery would provide no power.

**Solar Cells** A solar cell provides power upon illumination by photons. The photons create electron-hole pairs by promoting electrons into the conduction band from the

valence band. The electron-hole pairs are then separated into electrodes that have an affinity for either electrons or holes but not both. The affinity can be manifested by empty molecular orbitals in the electrode material that electrons and holes are drawn into, for example, in a polymer-inorganic photovoltaic cell. In typical p-n junction silicon solar cells, the photons create a chemical potential difference by altering the minority carrier concentration in the two junctions[18]. An often heard misconception is that the built-in electric field in a p-n junction solar cell is responsible for the separation which is not true. The separation is solely due to the chemical potential difference induced by the photons.

**Figure 1-2. The Chemical Potential Decreases as the Temperature is Raised for a 3-D Density of States.** As the temperature is raised, the Fermi distribution function changes from a step function at 0K to become more spread out. Since the DOS increases as $E^{1/2}$ for a 3-D system, the chemical potential must decrease as the temperature is raised above 0K. Otherwise, a situation with more electrons above the Fermi energy than empty states below would occur.

**Thermoelectrics** A thermoelectric device provides power when a temperature difference is applied to its ends. The chemical potential is sensitive to the temperature and, neglecting scattering effects, depends on the derivative of the density of states through the Sommerfeld approximation[4]:

$$\mu = E_F - \frac{\pi^2}{6}(kT)^2\left[\frac{\partial}{\partial E}\ln N(E)\right]_{E=E_F} + \dots \qquad \text{equation 4}$$

The total number of electrons must always be the same regardless of the temperature. The electron density is given by:

$$n = \int_0^\infty N(E)f(E)dE \quad \dots\dots\dots\text{equation 5}$$

For 3-D free electron systems, the DOS is proportional to $E^{1/2}$. The Fermi function $f(E)$ becomes more spread out as $T$ increases and since the electron density is a product of the DOS and Fermi function, the chemical potential must decrease (see Figure 1-2). Otherwise, one would encounter a situation with more electrons above the Fermi energy than empty states below[19]. Such a scenario would imply that electrons were added into the material as the temperature was increased. As long as the DOS is not constant near the Fermi energy, the temperature difference across a thermoelectric device will create a chemical potential difference that drives the diffusion of carriers (see Appendix 1.2).



**Figure 1-3. Thermoelectric Circuit.** Material A and Material B have different Seebeck Coefficients which causes a net current in the circuit. If both materials were the same, the net current would be zero.
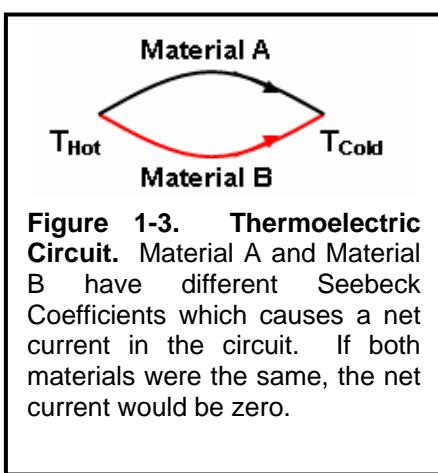
Similar to the situation in batteries, a thermoelectric circuit constructed with the same material for both electrodes, (Figure 1-3), would provide no power. The chemical potential difference is the same across both electrodes and leads to opposing currents that cancel.

### 1.3.4 Diffusion Thermopower in Semicondcutors

The expression for the thermopower must be modified for semiconductors since at the Fermi energy, the conductivity as expressed in equation 2 would be zero. Since, in general, the thermopower is the amount of net entropy transferred per charge carrier and since the entropy is just the heat carried by each charge carrier per unit temperature, we can write the thermopower as[4]:

$$S = \frac{Q}{T}\frac{1}{e} = \frac{E - \mu}{eT} \ldots\ldots\text{equation 6}$$

where $Q$ is the heat transported by the charge carriers, $T$ is the temperature, $e$ is the electronic charge, $E$ is the total energy, kinetic plus potential, that the carriers possess and $\mu$ is the chemical potential (Fermi energy). For lightly doped semiconductors, $E - \mu$ can be quite large so that semiconductors typically have larger thermopowers than metals[20]. Thermopower values are typically on the order of hundreds of microvolts. The thermopower in semiconductors is highly sensitive to the doping concentration since this controls $E - \mu$.

## 1.3.5    Phonon Drag Thermopower

A temperature gradient across a thermoelectric material causes the phonon system to be placed out of thermodynamic equilibrium. This is manifested in the form of a heat current where the phonon energy is propagated down the temperature gradient. If the phonon-carrier interaction is strong, the phonons may impart some of their energy to the diffusing carriers via momentum transfer[3]. At open circuit, the phonons effectively "drag" the carriers along the temperature gradient which causes an increase in the buildup of charge above that from diffusion of the carriers alone. The phonon drag and diffusion components can be treated independently and therefore the total thermopower can be written as $S = S_{diffusion} + S_{phonon-drag}$.

A classical treatment similar to that for the diffusion thermopower applies to the phonon system. The phonons can be treated as an isotropic gas that expands due to the temperature gradient. The expanding phonon gas exerts a pressure on the carriers that is biased towards the cold end due to the temperature gradient. The magnitude of the phonon drag thermopower depends on the strength of the interaction between the carriers

and phonons. Also, the phonon-carrier relaxation time should be shorter than the phonon-phonon relaxation time. That is because to enhance the thermoelectric voltage, the phonons should preferentially lose their energy to the carriers rather than to other phonon scattering events that arise from anharmonic coupling[3]. The fact that phonons may scatter with other phonons is the main reason why phonon drag in metals is only observed at low temperatures. At high temperatures phonons in metals lose most of their energy by scattering with other phonons. At very low temperatures, there are not enough populated phonon modes to significantly contribute to the thermopower. The thermopower due to phonon drag can be expressed as: $S_{ph} \approx \left( \dfrac{\tau_{ph}}{\mu T} \right)$ where $\tau_{ph}$ is the phonon-phonon relaxation time, $\mu$ is the carrier mobility and $T$ is the temperature[3]. Metals exhibit a thermopower that peaks between 4-50K due to phonon drag.



**Figure 1-4. Electron-Phonon Scattering.** The circle is a representation of the Fermi surface. An electron with initial momentum $\hbar k_1$ (blue arrow) is scattered by a phonon (purple arrow). Phonons with wavevector larger than $2k_F$ (where $k_F$ is the radius of the Fermi sphere) cannot scatter electrons. Therefore, only phonons with wavevectors < $2k_F$ can participate in phonon drag.
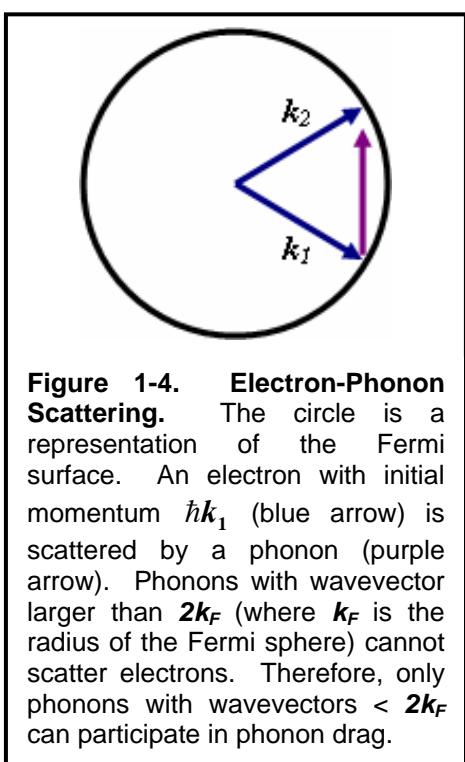
Phonons participating in phonon drag are typically low frequency because they scatter less often with other phonons than their high frequency counterparts. Only phonons with wave-vectors smaller than $2k_F$ can participate in phonon-carrier scattering (Figure 1-4). The Fermi surface in semiconductors is considerably smaller than metals since the carrier concentration is lower by ~3-4 orders of magnitude. Therefore, phonon drag dominates the thermopower in low doped semiconductors such as silicon since the low

frequency phonons are the dominant heat carriers[21].

## 1.4    Other Entropic Contributions to the Thermopower

In general any net transfer of entropy due to a temperature gradient will give rise to a thermoelectric voltage.  The thermopower can generally be expressed as:

$$S = \frac{S_{entropy}}{e} \qquad \ldots\ldots\ldots\text{equation 7}$$

where $S_{entropy}$ is the net entropy transferred in the presence of a temperature gradient and e is the electronic charge.  Magnetic systems such as $Na_xCo_2O_4$ can transfer spin entropy which gives rise to a thermopower[22].  According to equation 7, such systems exhibit a contribution from the spins given by:

$$S = \frac{\ln \Omega}{e} \qquad \ldots\ldots\ldots\text{equation 8}$$

where *ln $\Omega$* is the entropy from the spin degeneracy, $\Omega$.

## 1.5    Fundamental Principles—The Peltier Effect for Refrigeration

The Peltier effect describes the fact that a current through a thermoelectric material will



**Figure 1-5.  Peltier Effect for a Semiconductor/Metal Circuit.  a)** Only hot electrons at the right junction are able to enter the semiconductor's conduction band.  These electrons remove heat from the junction and subsequently cool it down.  Electrons from the semiconductor enter the metal at the left junction and thermalize near the metal's Fermi energy.  This causes the left junction to become hot.  **b)**  The current direction is reversed which causes the left junction to become cold and the right junction to become hot.  Reprinted from Ref. 23.

generate a temperature difference across its ends. The Peltier effect is the reverse of the Seebeck effect. The current, supplied by a power generator, exchanges heat at the junctions with the electrodes[23] (Figure 1-5). There the heat is either absorbed or released. This causes on junction to become cold and the other hot. The cold junction can be used as a refrigerator. The Peltier effect is reversible meaning that the cold and hot junctions can change polarity if the current direction is reversed.

The magnitude of the Peltier effect is given by the Peltier coefficient, $\Pi$, which is the ratio between the heat current and electric current. The Peltier coefficient is related to the Seebeck coefficient by the following relation: $\Pi = ST$.

## 1.6    References

1.  Seebeck, T.J., Magnetische polarisation der metalle und erze durck temperatur-differenz, *Abh. K. Akad. Wiss.*, **265** (1823).
2.  Peltier, J.C., Nouvelles experiences sur la caloriecete des courans electriques. *Ann. Chem.* **LVI**, 371-387 (1834).
3.  MacDonald, D.K.C., *Thermoelectricity: An Introduction to the Principles* (Wiley, New York, NY 1962).
4.  Ashcroft, N.W. & Mermin, N.D., *Solid State Physics* (Saunders College, FortWorth, TX 1976).
5.  Mahan, G., Sales, B. & Sharp, J. Thermoelectric materials: New approaches to an old problem. *Phys. Today* **50**, 42-47 (1997).
6.  Rowe, D.M. Ed., *CRC Handbook of Thermoelectrics* (CRC Press, Boca Raton, Fl 1995).
7.  Ioffe, A.F., *Semiconductor Thermoelements and Semiconductor Cooling* (Infosearch, London, UK 1957).
8.  Majumdar, A. Enhanced: Thermoelectricity in semiconductor nanostructures. *Science* **303**, 777-778 (2004).
9.  Chen, G. *et al*. Recent developments in thermoelectric materials. *Int. Mater. Rev.* **48**, 45-66 (2003).
10. Venkatasubramanian, R. *et al*. Thin-film thermoelectric devices with high room-temperature figures of merit. *Nature* **413**, 597-602 (2001).
11. Harman, T.C. *et al*. Quantum dot superlattice thermoelectric materials and devices. *Science* **297**, 2229-2232 (2002).
12. Hsu, K.F. *et al*. Cubic $AgPb_mSbTe_{2+m}$: Bulk thermoelectric materials with high figure of merit. *Science* **303**, 818-821 (2004).

13. Li, D. *et al*. Thermal conductivity of individual silicon nanowires. *Applied Physics Letters* **83**, 2934-2936 (2003).

14. Humphrey, T.E. & Linke, H. Reversible thermoelectric nanomaterials. *Physical Review Letters* **94**, 096601 (2005).

15. Callen, H.B., *Thermodynamics* (Wiley, New York, NY 1985).

16. Mott, N.F. & Jones, N., *The Theory of the Properties of Metals and Alloys* (Dover, New York, NY 1958).

17. Kittel, C. & Kroemer, H., *Thermal Physics* (W.H. Freeman, New York, NY 1980).

18. Wurfel, P., *Physics of Solar Cells: From Principles to New Concepts* (Wiley, Weinheim, Germany 2005).

19. Schroeder, D.V., *An Introduction to Thermal Physics* (Addison Wesley Longman, San Francisco, CA 2000).

20. Sze, S.M., *Physics of Semiconductor Devices* (Wiley, New York, NY 1981).

21. Boukai, A.I. *et al*. Silicon Nanowires as Efficient Silicon Nanowires. *Nature*, (accepted 11/2/07).

22. Wang, Y *et al*. Spin entropy as the likely source of enhanced thermopower in $Na_xCo_2O_4$. *Nature* **423,** 425 (2003).

23. Winder, E.J. *et al*. Thermoelectric Devices: Solid- State Refrigerators and Electrical Generators in the Classroom. *Journal of Chemical Education* **73**, 940 (1996).

**Appendix 1.1**

## Derivation of the Efficiency for a Thermoelectric Device
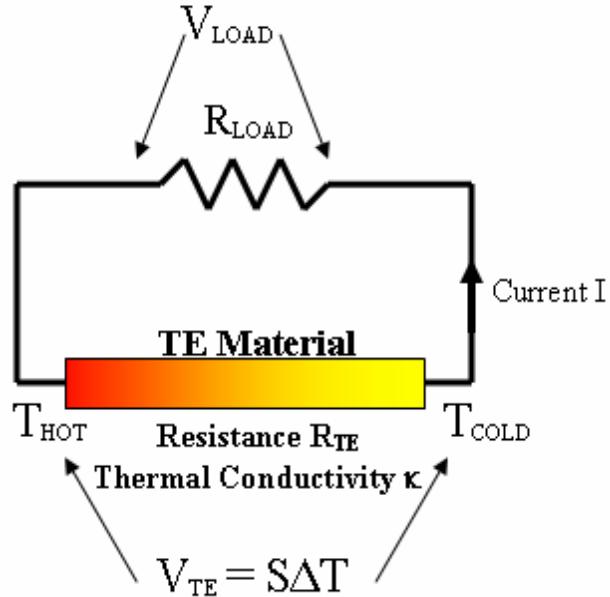


1.  The efficiency $\eta$ is given by: $\eta = \dfrac{W}{Q}$ where W is the work output of the TE generator and Q is the heat input into the hot reservoir.

2.  The useful work that can be extracted from the TE generator is equal to the power dissipated in the load resistor. $W = IV_{LOAD}$ .

3.  From voltage division: $V_{LOAD} = V_{TE}\left(\dfrac{R_{LOAD}}{R_{LOAD} + R_{TE}}\right) = S\Delta T\left(\dfrac{R_{LOAD}}{R_{LOAD} + R_{TE}}\right)$.

4.  The current is: $I = \dfrac{S\Delta T}{R_{LOAD} + R_{TE}}$ .

5.  Therefore: $W = S^2 \Delta T^2 \left[\dfrac{R_{LOAD}}{\left(R_{LOAD} + R_{TE}\right)^2}\right]$ .

6.  The heat input into the TE generator can be broken up into three terms: a) the heat that originates at the hot junction via standard thermal conduction (phonons and carriers) $Q_1 = \kappa\Delta T$ , this heat is transferred to the cold junction, b) the heat that

originates at the hot junction by the Peltier effect due to the current

$$Q_2 = IST_{HOT} = \frac{S^2 \Delta T T_{HOT}}{(R_{LOAD} + R_{TE})},$$ this heat is also transferred to the cold junction,

and c) the heat that originates via joule heating in the material half of which continues to the cold junction and half of which is transferred back to the hot junction; the latter is what we are concerned about since Q is the heat input at the

hot junction $$Q_3 = -\frac{1}{2} I^2 R_{TE} = -\frac{1}{2} \frac{S^2 \Delta T^2 R_{TE}}{(R_{LOAD} + R_{TE})^2}.$$

7. Therefore combining the terms:

$$Q = \kappa \Delta T + \frac{S^2 \Delta T T_{HOT}}{(R_{LOAD} + R_{TE})} - \frac{1}{2} \frac{S^2 \Delta T^2 R_{TE}}{(R_{LOAD} + R_{TE})^2}.$$

8. So the efficiency becomes: $$\eta = \frac{S^2 \Delta T^2 \left[ \dfrac{R_{LOAD}}{(R_{LOAD} + R_{TE})^2} \right]}{\kappa \Delta T + \dfrac{S^2 \Delta T T_{HOT}}{(R_{LOAD} + R_{TE})} - \dfrac{1}{2} \dfrac{S^2 \Delta T^2 R_{TE}}{(R_{LOAD} + R_{TE})^2}}.$$

9. We now want to know how to maximize the efficiency. This will depend on the ratio of the load resistance to the TE resistance. We will define this ratio as

$m = \dfrac{R_{LOAD}}{R_{TE}}$. We know from standard electronics that maximum power transfer

occurs when $R_{LOAD} = R_{TE}$ so that $m = 1$. But what value will m take for maximum efficiency?

10. We can rewrite the efficiency as: $$\eta = \frac{S^2 \Delta T^2 \dfrac{m}{R_{TE}(m+1)^2}}{\kappa \Delta T + \dfrac{S^2 \Delta T T_{HOT}}{R_{TE}(m+1)} - \dfrac{1}{2} \dfrac{S^2 \Delta T^2}{R_{TE}(m+1)^2}}.$$

11. Reducing this equation further by pulling out the term: $\dfrac{S^2 \Delta T T_{HOT}}{R_{TE}(m+1)}$ out of the

denominator, we have: $\eta = \dfrac{\Delta T}{T_{HOT}}\left[\dfrac{\dfrac{m}{(m+1)}}{1 + \dfrac{\kappa R_{TE}}{S^2}\dfrac{(m+1)}{T_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{1}{(m+1)}}\right]$

12. Now we are in the position to maximize $\eta$.

13. $\dfrac{\partial \eta}{\partial m} = 0 =$

$$\dfrac{\Delta T}{T_{HOT}}\left\{\dfrac{\left[\dfrac{1}{(m+1)} - \dfrac{m}{(m+1)^2}\right]\left[1 + \dfrac{\kappa R_{TE}}{S^2}\dfrac{(m+1)}{T_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{1}{(m+1)}\right] - \left(\dfrac{m}{m+1}\right)\left[\dfrac{\kappa R_{TE}}{S^2 T_{HOT}} + \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}(m+1)^2}\right]}{\left[1 + \dfrac{\kappa R_{TE}}{S^2}\dfrac{(m+1)}{T_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{1}{(m+1)}\right]^2}\right\}$$

14. Defining $Z = \dfrac{S^2}{\kappa R_{TE}}$, we have:

$$0 = \dfrac{1}{(m+1)} - \dfrac{m}{(m+1)^2} + \dfrac{1}{ZT_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}(m+1)^2} - \dfrac{1}{ZT_{HOT}}\dfrac{m}{(m+1)} + \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{m}{(m+1)^3} - \dfrac{m}{(m+1)}\dfrac{1}{ZT_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{m}{(m+1)}$$

15. $0 = \dfrac{1}{(m+1)} - \dfrac{m}{(m+1)^2} + \dfrac{1}{ZT_{HOT}} - \dfrac{2m}{ZT_{HOT}(m+1)} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}(m+1)^2}$

16. Multiplying through by $(m+1)^2$, we have:

$$m + 1 - m + \dfrac{(m+1)^2}{ZT_{HOT}} - \dfrac{2m(m+1)}{ZT_{HOT}} - \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}}\dfrac{1}{(m+1)^2} = 0$$

17. $\dfrac{(m+1)^2}{ZT_{HOT}} - \dfrac{2m(m+1)}{ZT_{HOT}} = \dfrac{1}{2}\dfrac{\Delta T}{T_{HOT}} - 1$

18. Multiplying through by $ZT_{HOT}$, we have:

$$m^2 + 2m + 1 - 2m^2 - 2m = \frac{1}{2}Z\Delta T - ZT_{HOT}$$

19. $m^2 = 1 + \frac{Z}{2}\left(2T_{HOT} - T_{HOT} + T_{COLD}\right)$

20. Finally: $m = \sqrt{1 + \frac{Z}{2}\left(T_{HOT} + T_{COLD}\right)} = \sqrt{1 + Z\overline{\overline{T}}}$ where $\dfrac{T_{HOT} + T_{COLD}}{2} = \overline{T}$

21. Substituting back into the equation for the efficiency, we have:

$$\eta_{max} = \frac{\Delta T}{T_{HOT}} \frac{\sqrt{1 + Z\overline{\overline{T}}} - 1}{\sqrt{1 + Z\overline{\overline{T}}} + \dfrac{T_{COLD}}{T_{HOT}}}$$

22. We can see that as $Z\overline{T} \to \infty$, $\quad \eta_{max} = \dfrac{\Delta T}{T_{HOT}}$ (the Carnot efficiency)

**Appendix 1.2**

## Derivation of the Thermopower Dependence
## on the Density of States

Consider a thermoelectric material with density of states *N(E)* and temperature difference

$\Delta T = T_{Hot} - T_{Cold}$ across its ends.  Using the first two terms of the Sommerfeld expansion,

the chemical potential can be approximated by:

$$\mu = E_F - \frac{\pi^2}{6}(kT)^2\left[\frac{\partial}{\partial E}\ln N(E)\right]_{E=E_F}$$

The chemical potential at the cold end of the thermoelectric is given by:

$$\mu = E_F - \frac{\pi^2}{6}(kT_{Cold})^2\left[\frac{\partial}{\partial E}\ln N(E)\right]_{E=E_F}$$

The chemical potential at the hot end of the thermoelectric is given by:

$$\mu = E_F - \frac{\pi^2}{6}[k(T_{Cold}+\delta T)]^2\left[\frac{\partial}{\partial E}\ln N(E)\right]_{E=E_F}$$

where δT is << 1.  Therefore, there is a chemical potential difference between the hot and

cold ends of the thermoelectric material:

$$\Delta\mu = \frac{\pi^2}{6}k^2(2T_{Cold}\delta T)\left[\frac{\partial}{\partial E}\ln N(E)\right]_{E=E_F}$$

The carriers diffuse because the temperature difference creates a chemical potential

difference.  The diffusion of the charged carriers creates an electric field that opposes

further diffusion.  Eventually, a dynamic equilibrium exists and the chemical potential

difference equals the voltage difference that opposes further diffusion:

$$\Delta\mu = e\Delta V = \frac{\pi^2}{6} k^2 (2T_{Cold}\delta T) \left[ \frac{\partial}{\partial E} \ln N(E) \right]_{E=E_F}$$

Therefore, the thermopower is:

$$S = \frac{\Delta V}{\delta T} = \frac{\pi^2}{3} \frac{k^2}{e} T_{Cold} \left[ \frac{\partial}{\partial E} \ln N(E) \right]_{E=E_F}$$

This is identical to the Mott formula derived using Boltzmann's transport equation except

that the dependence of scattering has not been taken into account.

# Chapter 2

# Fabrication and Measurements of Bismuth and Silicon Thermoelectric Devices

**The results presented in this chapter and chapters 3 and 4 are largely extracted from A. Boukai, *et al*. Advanced Materials 18, 864–869 (2006) and A.I. Boukai *et al*. Nature 451, 168–171 (2008).**

## 2.1    Bismuth and Silicon Nanowires for Use as Thermoelectrics

Bismuth has a small effective mass ($0.001\boldsymbol{m_0}$ along the trigonal direction), large thermoelectric power (50 µV/K to 100 µV/K), and small thermal conductivity (8 W·m$^{-1}$·K$^{-1}$), all of which help increase $\boldsymbol{ZT}$[1]. For bulk Bi, $\boldsymbol{ZT}$ is limited because the conduction and valence bands indirectly overlap (~38 meV), which causes Bi to be a compensated semimetal,[2] meaning that the contributions of holes and electrons to $\boldsymbol{S}$ cancel each other. However, a semimetal-to-semiconductor transition has been predicted for Bi nanowires below a critical size (~50 nm depending on the crystal orientation) due to the quantum confinement effect[1,3].  Below this size, Bi nanowires can be doped either n- or p-type which causes $\boldsymbol{S}$ to be significantly enhanced for Bi nanowires (NWs) with optimized position of the Fermi energy. 1-D semiconductors are also characterized by a sharply peaked density of states[4,5]. Experimental evidence of the enhancement of $\boldsymbol{S}$ has been reported in Bi nanocomposites[6]. A Bi nanowire will also have a reduced $\kappa$ due to phonon scattering off the sidewalls which helps increase $\boldsymbol{ZT}$[1,7].

Bulk Si is a poor thermoelectric due to its large thermal conductivity, 150 $W \cdot m^{-1} \cdot K^{-1}$, at room temperature. Si nanowires have been shown to have dramatically reduced thermal conductivities. Therefore, if the electrical conductivity and thermopower of a Si nanowire can be maintained similar to the bulk, the Si nanowires have the potential to be efficient thermoelectrics.

## 2.2    Fabrication of Bismuth Nanowires

Owing to its chemical instability and low melting point (~271 °C), fabrication and electric measurements of Bi nanowires are challenging. Embedded nanowire arrays have been prepared in porous templates by various methods[8-10]. Transport properties of such embedded NWs are limited to two-point resistance measurements,[9,10,11-16] with the numbers of wires measured being indeterminate[17,18]. Fewer measurements on the thermoelectric power of Bi nanowire arrays have been reported[15,16]. In addition, the very small thicknesses of the embedded NW samples leads to an underestimate of $S$ because the thermocouples used to measure the temperature are comparable to the sample thickness[17].

We have overcome several of these limitations by coupling a device architecture used by Small *et. al.*[19] and Llaguno *et. al.*[20] for nanotube studies with nanofabrication steps that are compatible with the material limitations of bismuth. This resulting device provides a reliable four-point electrical conductivity measurement, allows for accurate measurements of thermoelectric power, and permits quantitation of electric and magnetic field effects on the transport of individual Bi NW.

Figure 2-1(a) shows one of our devices. Thin Pt/Ti (7 nm/3 nm) electrodes and Pt/Ti (50 nm/10 nm) heaters are fabricated onto a thermal $SiO_2$ coated (500 nm thick) degenerately doped Si substrate by electron beam lithography (EBL). As the last step, ~2 μm long NWs are defined across the Pt electrodes by EBL. 15 nm of bismuth (99.999%, Puratronic, Alfa Aesar) is e-beam evaporated at room temperature at 0.5 Å/s with a base pressure of $< 5 \times 10^{-7}$ Torr. A 28 nm wide Bi NW is
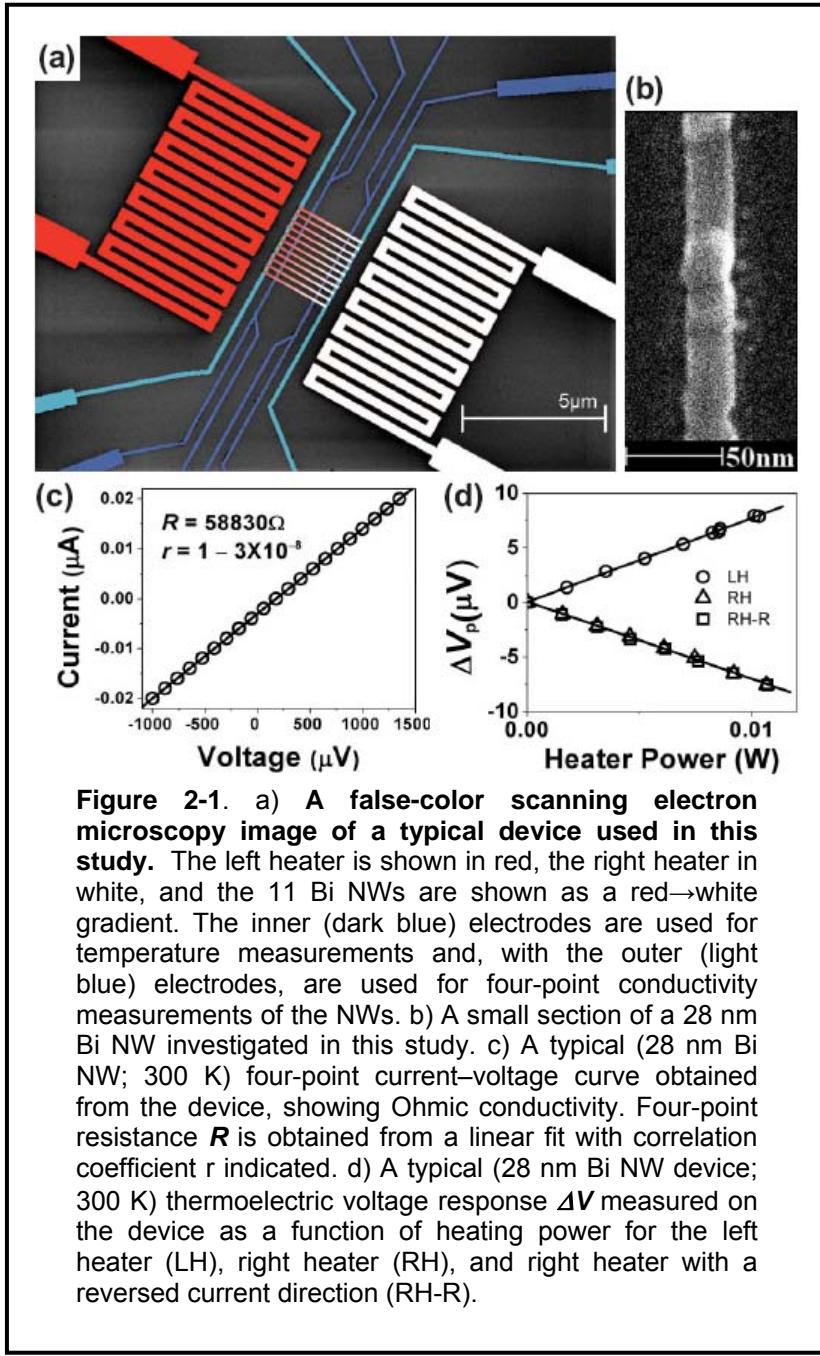


**Figure 2-1**. a) **A false-color scanning electron microscopy image of a typical device used in this study.** The left heater is shown in red, the right heater in white, and the 11 Bi NWs are shown as a red→white gradient. The inner (dark blue) electrodes are used for temperature measurements and, with the outer (light blue) electrodes, are used for four-point conductivity measurements of the NWs. b) A small section of a 28 nm Bi NW investigated in this study. c) A typical (28 nm Bi NW; 300 K) four-point current–voltage curve obtained from the device, showing Ohmic conductivity. Four-point resistance **R** is obtained from a linear fit with correlation coefficient r indicated. d) A typical (28 nm Bi NW device; 300 K) thermoelectric voltage response **ΔV** measured on the device as a function of heating power for the left heater (LH), right heater (RH), and right heater with a reversed current direction (RH-R).

shown in Figure 2-1(b). The NWs are polycrystalline with an in-plane grain size comparable to the wire width. XRD results on a 15 nm thick film show the (003) and (006) peaks of rhombohedral Bi, with grains aligning to the (001) trigonal orientation

normal to the substrate, but likely with arbitrary in-plane orientations[21]. Dimensions are measured with a scanning electron microscope (SEM) and atomic force microscopy (AFM).

The Bi NWs are extremely fragile due to their small size, low melting point, and chemical instability. They are easily broken by ambient static charges and oxidation typically causes a 1% increase of the measured four-point resistance for every hour of air exposure. Thus, the NWs are fabricated just prior to wire bonding and chip mounting. During these last steps, the device contacts are shorted through a guard ring that is disconnected right before the device is loaded into a helium-sealed thermostat. With this approach, the NWs remain intact and no discernible increase of resistance is found during the course of the few-week measurement period.

## 2.3    Thermoelectric Measurements on Bismuth Nanowires

The outer electrodes pass a current through the NWs, while the inner two sets serve as voltage probes for the four-point measurements and the thermoelectric voltage measurements. A current passed through a heater locally heats the $SiO_2$, resulting in a temperature gradient (and thermoelectric voltage) along the Bi NWs. The thermoelectric voltage is detected from the inner two sets of electrodes with a Keithley 2182 nanovoltmeter (impedance $>10G\Omega$), as a function of the heater power [Figure 2-1(d)]. The temperature difference between these inner electrodes is calibrated to the heater power by monitoring the respective changes in their four-point resistance when changing the heater power or system temperature (within 4K). The calibration is done on every

device for each temperature point that the thermoelectric power is measured. The highly doped Si substrate underneath $SiO_2$ serves as a gate electrode for electric field effect measurements.

Care was taken to eliminate possible experimental artifacts. For four-point resistance measurements, *I-V* curves [Figure 2-1(c)] are always obtained with good linearity, confirming ohmic conductivity of the NWs. The resistance is obtained from the slope rather than from a single-point *I-V* calculation to avoid possible influences from offset voltages. For thermoelectric measurements, either of the heaters is turned on right after re-zeroing the nanovoltmeter, and the voltage change $\Delta V$ is recorded as the thermoelectric voltage after the system has reached thermal equilibrium. $\Delta V$ originates from the temperature gradient caused by local heating based on the observations that $\Delta V$ increases linearly with the increase of heater power (square of the current), and that a reversed direction of current in the heater (which produces the same joule heat) doesn't change $\Delta V$. Finally, passing a current through the heater on the other side of the NWs produces a comparable $\Delta V$, but with opposite sign [Figure 2-1(d)]. The measured thermopower contains the thermopower contribution from the electrical leads (Appendix 2.1).

## 2.4　Fabrication of Silicon Nanowires

The devices used were fabricated from intrinsic silicon-on-insulator (SOI) substrates (Ibis Technology) with a buried oxide thickness of 150nm. The top Si(100) epilayer began with a thickness of 400nm and was thinned through repeated thermal oxidations and wet

chemical etching steps until the desired Si layer thickness (20nm or 35nm) was obtained. AFM analysis revealed a smooth surface with an rms roughness of ~3.6 Å. Next, the Si epilayer was doped using a boron containing spin-on dopant (Filmtronics Boron-A) (21). After spin coating, the substrates were baked (250 ºC; 5 min) and annealed under flowing $N_2$ using a rapid thermal processor at temperatures corresponding to the desired doping concentration by allowing the boron to diffuse into the silicon. Four-point probe conductivity measurements determined the p-type doping concentration. We have previously reported that the doping profile through the thickness of the epilayer drops rapidly from the top surface (Figure 2-2)[22].

There are multiple ways to prepare Si NWs, including materials methods for bulk production. We wanted Si NWs in which the dimensions, impurity doping levels, crystallographic nature, etc., were all quantifiable and precisely controlled. Thus we utilized the Superlattice Nanowire Pattern transfer (SNAP) method[23] that translates the atomic control over the layer thickness of a superlattice into control over the width and spacing of NWs. Si NWs made via SNAP inherit their impurity dopant concentrations directly from the single-crystal Si epilayers of the silicon-on-insulator (SOI) substrates from which they are fabricated. These epilayers were 20 or 35 nm thick Si(100) films on 150 nm of $SiO_2$, and were p-type impurity (Boron) doped using diffusion based doping[22]. Four-point probe conductivity measurements of the SOI films were utilized to extract dopant concentrations. We prepared NW arrays several micrometers long, with lateral dimensions of 10×20, 20×20, and 520×35 (width×thickness in nm). This last one approximates the bulk and, in fact, measurements on the sample obtained bulk values for $S$, $\sigma$, and $\kappa$.

After the diffusion doping process, the substrates were used to fabricate Si NWs using the SNAP process[23]. The PMMA/epoxy used to hold the superlattice onto the SOI consisted of .37g of 6% by weight PMMA dissolved in 20 mL of chlorobenzene with the addition of 20 drops part A epoxide and 2 drops part B hardener (Epoxy Bond 110, Allied High Tech, Rancho Dominguez, CA). The depth of the NWs was nominally 20nm (the SOI substrate thickness). However, the "electrical" depth, the distance at which the doping concentration decreased by an order of magnitude or more, was 10-15 nm (Figure 2-2). SEM and AFM were used to
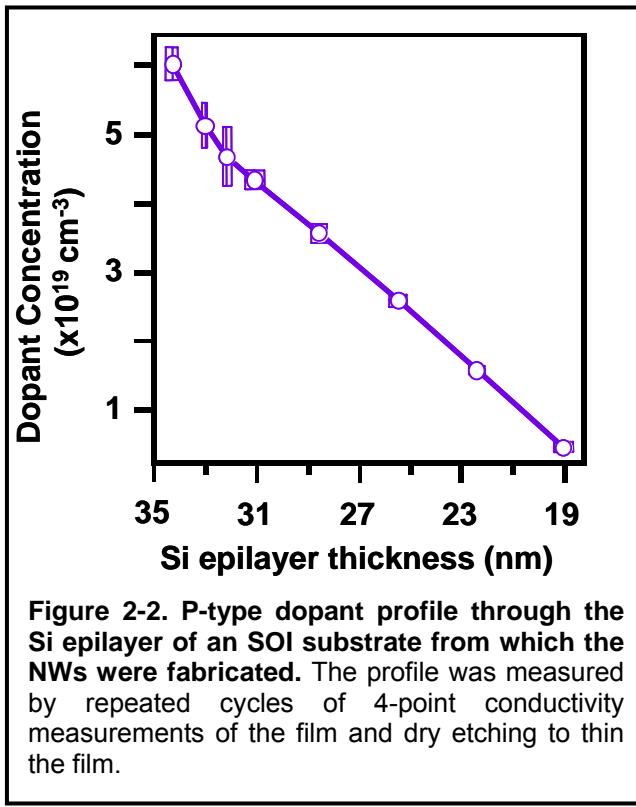
**Figure 2-2. P-type dopant profile through the Si epilayer of an SOI substrate from which the NWs were fabricated.** The profile was measured by repeated cycles of 4-point conductivity measurements of the film and dry etching to thin the film.

determine the dimensions of our Si NWs. Devices contained from 10 to 400 NWs connected in parallel. For all measurements on the 10 nm wide wires other than thermal conductivity, monolithic contacts were used in order to minimize the contact resistance (Figure 2-3.). The use of monolithic contacts means that the NWs constitute a bridge between two micrometer-sized pads, with the NWs and pads all part of the same single crystal silicon structure. This method effectively removes contact resistance.

Measurement of the three thermoelectric parameters, $S$, $\sigma$, and $\kappa$ was accomplished using a device architecture similar to those of other published reports[19,24-25] (Figure 2-4). Heaters and electrodes were fabricated using e-beam lithography (EBL)

and e-beam evaporation of metal. The heaters and electrodes were both composed of titanium/platinum and were 20nm/100nm thick and 20nm/160nm thick respectively. Prior to metallization of the electrodes, the Si NWs were dipped in a buffered HF solution for 3 seconds to remove any native oxide. After electrode metallization, the device was annealed in forming gas inside the RTP ($475^{o}$C; 5 min). The electrodes were then contacted by large gold pads (10nm/220nm Ti/Au) that were defined by photolithography.
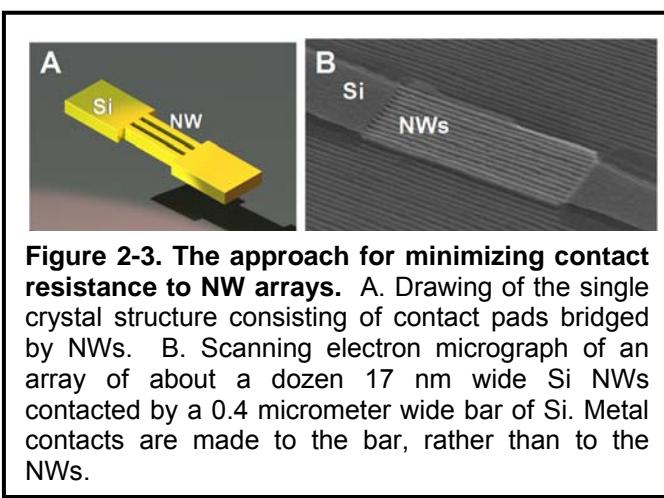


**Figure 2-3. The approach for minimizing contact resistance to NW arrays.** A. Drawing of the single crystal structure consisting of contact pads bridged by NWs. B. Scanning electron micrograph of an array of about a dozen 17 nm wide Si NWs contacted by a 0.4 micrometer wide bar of Si. Metal contacts are made to the bar, rather than to the NWs.

We used EBL and several etching steps to suspend the Si NWs. Polymethyl methacrylate (PMMA), an EBL resist material, is spun coated onto the surface of a chip containing a fully fabricated device. EBL is used to define an opening in PMMA around the active area of the device. A small rectangular island of PMMA remains on top of the NWs, electrodes and heaters to protect them during subsequent etching steps. Next, we used a $CF_4$ plasma etch, 20 minutes at 40W and 20 mTorr, to remove the $SiO_2$ around the active region. A 1.5 minute $XeF_2$ etch removes the exposed silicon and undercuts the silicon handle layer underneath the active region. During this step, the NWs are protected by the thick rectangular island of PMMA on top and the 150nm thick oxide underneath. The PMMA is then removed using an $O_2$ plasma (5 min; 40W; 20 mTorr). At the end, the Si NWs remain anchored on top of a thin rectangular island of 150nm thick $SiO_2$.

Removal of the oxide island would cause the NWs to collapse, and so, as described below, measurements of $\kappa$ were carried out as a differential measurement – i.e. a measurement of the NWs + oxide island followed by a measurement of the oxide island. Since this is a measurement of the heat flow only through the NWs, the oxide acts as a parasitic path for this heat flow that must be experimentally quantified.

Electrodes contacting the Si NWs were initially shorted together to minimize device damage through static charge. When the measurement begins, the shorts were disconnected. The Si substrate is mounted onto a thermally conductive chip carrier with the gold pads wire-bonded to that same chip carrier. The chip carrier is then mounted onto the cold finger of a vacuum sealed liquid nitrogen cryostat (Janis Research VPF-475), which is evacuated to $\leq 10^{-6}$ Torr for the measurements. A temperature controller (Lakeshore Model 331) allows global control of the cryostat temperature from 350K to 77K with a precision of .01K.
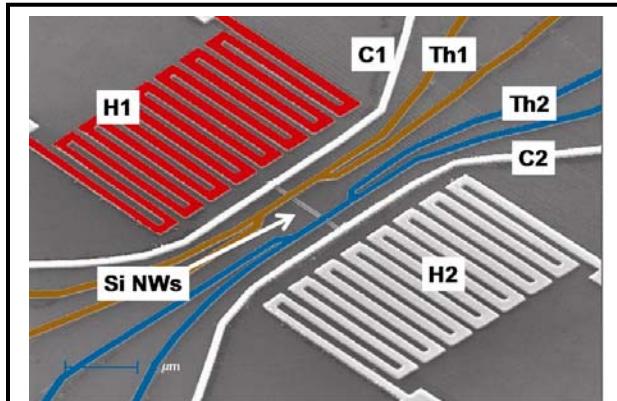


**Figure 2-4. False color SEM image of a (non-suspended) thermoelectric device platform used in this study.** The heaters (H1 and H2) are shown, one in red to indicate that Joule heating is done using just one of the heaters. C1, C2, Th1 and Th2 comprise the 4-point electrical contacts to the Si NWs, with C1 and C2 utilized as the current source and drain for those electrical conductivity measurements. Th1 and Th2 also serve as resistive thermometers (note that these two electrodes are themselves 4-point devices).

## 2.5    Thermoelectric Measurements on Silicon Nanowires

Determining the thermopower, $\dfrac{\Delta V}{\Delta T}$, requires measurements of the thermoelectric voltage

$\Delta V$ and the temperature difference $\Delta T$ along the Si NWs. A switching matrix (Keithley 707A) allows reconfigurable electrical connections from device to instrumentation for the various thermoelectric measurements.

We first create a temperature gradient along the Si NWs using a DC current through one heater (Joule heating) (Figure 2-4). We then measure the resultant thermoelectric voltage, $\Delta V$ (TEV). A Keithley 2182A nano-voltmeter measures the TEV between the inner electrodes (Th1 and Th2). In order to verify that the measured voltage is truly a result of the temperature gradient, the measurement is repeated using the second heater situated on the opposite side of the NW circuit. This should yield a voltage with similar magnitude but opposite polarity. A second check is to reverse the current polarity in the heater – this should not alter the magnitude or the polarity of the thermoelectric voltage. Figure 2-5A verifies that the measured voltage is indeed due to the temperature gradient.

Two separate measurements yield $\Delta T$ by using the inner electrodes as resistive thermometers. A resistive thermometer responds to temperature changes with a proportionate change in resistance. In the first measurement, the resistance of the Ti/Pt electrodes is continuously recorded by a four-point probe measurement as a function of heater power with one heater on. As the power dissipated in the heater slowly ramps stepwise, four lock-in amplifiers (Stanford Research Systems SR-830) simultaneously monitor the resistance of both thermometers. The first pair of lock-in units measures the

AC current and voltage at 913 Hz for one thermometer.  Another pair of lock-in units

performs the same functions for the other thermometer at 921 Hz.  We obtain the

resistances, $R$, of each thermometer as a function of heater power, $Q$ (Figure 2-5B).

In the second measurement, we ramp the cryostat temperature stepwise by 1K

increments while simultaneously measuring the thermometer's resistance.  We obtain a

linear relationship between the change in resistance and the cryostat temperature (Figure

2-5C).  This relates our thermometer's resistance change to a temperature difference,

$\Delta T$ , through the relation:

$$\Delta T = W \bullet \left( \frac{dR_{Left}/dQ}{dR_{Left}/dT} - \frac{dR_{Right}/dQ}{dR_{Right}/dT} \right)$$

where $R_{Left}$ and $R_{Right}$ are the resistances of the left and right thermometers respectively.

The slope of a linear fit to the TEV vs $\Delta T$  data provides the value of the thermopower.

This process is repeated to calculate the thermopower from 300K down to 100K.
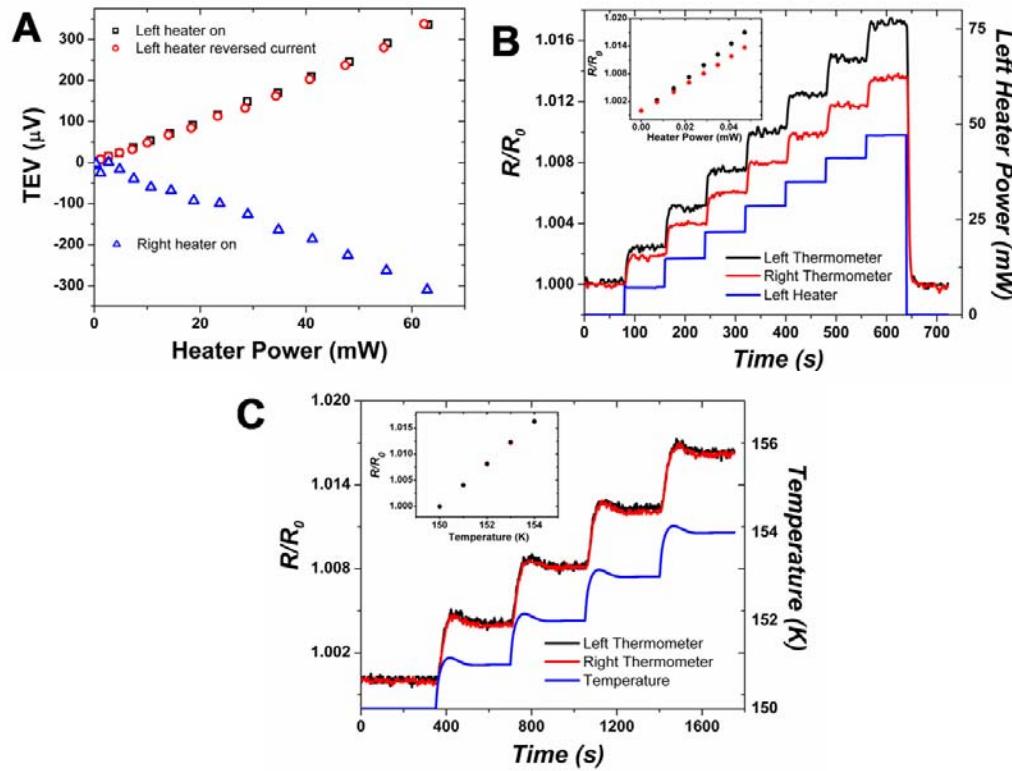
**Figure 2-5. Measurement of the thermoelectric power of Si NWs. A.** Plot of the thermoelectric voltage (TEV), which is a linear function of heater power at 150K. Each point represents the average of 6 measurements. With the left heater on, the TEV has no dependence on the heating current direction. With the right heater on, the TEV has the opposite polarity. **B.** Thermometer calibration. The graph shows the simultaneously measured resistance of both left and right thermometers (black and red trace, left y-axis) in response to the stepwise increase in heater power (blue trace, right y-axis) at 150K. The inset shows the linear relationship of the normalized thermometer resistance as a function of heater power. The slope of the linear fit to these two lines provides $dR_{Left}/dQ$ and $dR_{Right}/dQ$. Each point in the inset corresponds to an average of 20 points collected from the time plot at each plateau. **C.** Graph of the simultaneously measured resistance of both thermometers (black and red trace, left y-axis) as a function of global temperature in the cryostat (blue trace right y-axis). Inset shows a linear relationship of the normalized thermometer's resistance to the cryostat temperature. The slope of the linear fit to these two lines provides $dR_{Left}/dT$ and $dR_{Right}/dT$. Each point in the inset represents an average of 150 points collected in the time plot at each plateau.
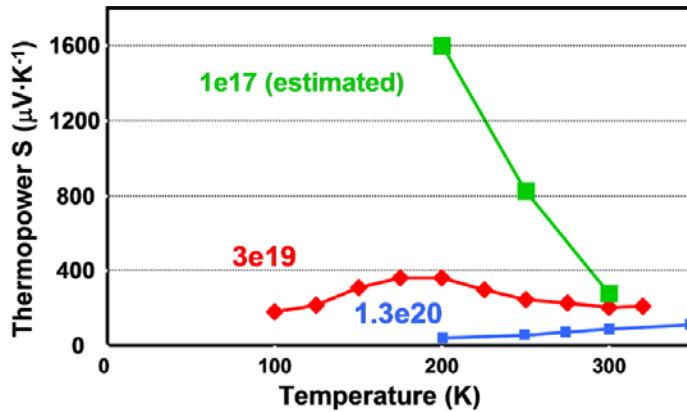
**Figure 2-6. Thermopower measurements for various 20nm x 20nm Si NWs.** The doping level of the low-doped sample (green) was estimated from NW conductivity measurements, instead of 4-point conductivity measurements of the SOI Si film from which the NWs were fabricated. As the doping is decreased, the measured **S** increases dramatically. This is consistent with phonon drag observed in lightly doped semiconductors. At lower doping, the Fermi surface is smaller and long wavelength acoustic phonons absorb the electron momentum. Since long wavelength phonons have longer lifetimes, the phonon drag contribution to the thermopower increases as the doping decreases.

## 2.5.1    Thermal Conductivity:  Measurement and Data

Techniques utilized to measure the thermal conductivity, $\kappa$, were developed by Majumdar's group[26,27]. The general expression for the thermal conductance of any system is given by the expression:  $Q=K\Delta T$, where $Q$ is the heat transferred across the NWs, $\Delta T$ is the temperature difference and $K$ is the thermal conductance ($K$ is related to $\kappa$ through geometrical factors).  Measurement of $\kappa$ is not as straightforward since parasitic pathways for the heat flow, $Q$, must be accounted for. Therefore, we determine the parasitic heat flow through the beams that support the suspended structure. A direct measurement of the power dissipated in the heater, $Q_{Heater}$, is possible by using the 4-point probe electrodes attached to the heater's windings since the power dissipated in the heater is equal to the product of its current and voltage drop. A lock-in measurement similar to that described for measurement of the thermopower allows us to measure $\Delta T$.

Also, the thermometers are in intimate contact with the ends of the NWs, eliminating any contributions from contact thermal resistances. The resistance of the thermometers is typically two orders of magnitude smaller than the resistance of the NWs. Taking into account the thermal conductance of the beams, $K_b$, the thermal conductance of the sample, $K_s$, is given by (17):

$$K_s = K_b \frac{T_C}{\Delta T} \ldots\ldots\ldots \text{eq. 1}$$

where $T_C$ is the temperature increase at the cold junction and $\Delta T$ is the temperature difference between the hot and cold junctions. The thermal conductance of the beams is given by:

$$K_b = \frac{Q}{T_H + T_C} \ldots\ldots\ldots \text{eq. 2}$$

where $Q$ is the power dissipated in the heater and $T_H$ is the temperature increase of the hot junction.

Our suspended structure is a combination of the $SiO_2$ island that supported the NWs and electrodes, and the Si NWs. To estimate the contribution of the $SiO_2$ to the measured $K_{total}$, we selectively



**Figure 2-7.** A) Thermal conductivity measurements of various Si NWs versus Temperature. Error bars are at 95% confidence. B) Thermal conductivity of a bulk Si device.

remove the NWs with a $XeF_2$ etc[28] and then measure $K_{SiO2}$. The thermometers were

separately calibrated during each thermal conductance measurement – both before and after the Si NWs were removed. $K_{SiO2}$ is subtracted from $K_{total}$, giving $K_{NWs}$. Finally, $K_{NWs}$ is converted to the $\kappa$ from dimensions and number of NWs (Figure 2-7). The $\kappa_{bulkSi}$ ~150 W m$^{-1}$K$^{-1}$ at 300K. $\kappa$ of thin films of Si has been found to be reduced by 50-70% over the bulk, depending upon the thickness, doping, and roughness of the film[29]. Our measured value of ~100 W m$^{-1}$K$^{-1}$ at 300K is somewhat higher than other reported values. This may arise from factors such as the smoothness of our epilayers or the doping profile through the thickness of our films. Our measured values are consistent with published reports from Majumdar's and Yang's groups[27]. Also, the change in the thermal conductance of the metallic beams before and after XeF$_2$ is much less than 1%. This indicates that the difference between the two values of the thermal conductance arises solely from removal of the NWs.

## 2.5.2    Electrical Conductivity:  Measurement and Data

The electrical conductivity of our NWs was measured using a four-point probe technique that removes any contribution from the contacts. A Keithley 2400 provides a current source through the outer electrodes and the nano-voltmeter measures the voltage drop between the two inner electrodes. The conductance value is calculated from the slope of a linear fit



**Figure 2-8.** **Representative electrical conductivity data for Si NWs and microwires**; p-type doping levels are indicated. All NWs are 20nm in height.
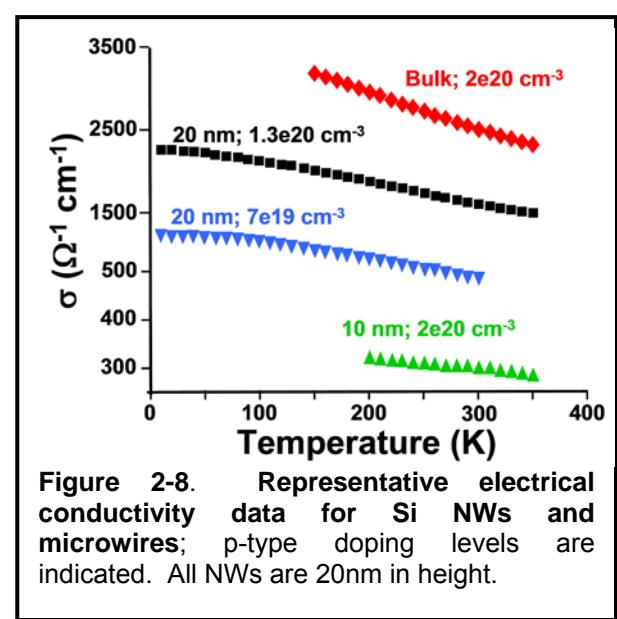
of the I-V data. The conductance is then converted to an electrical conductivity using the dimensions & numbers of the NWs (Figure 2-8).

The measured thermopower includes the contribution from the oxide substrates. We show that the contribution from the oxide thermopower $S_{oxide}$ is negligible because the oxide resistance is much larger than the NWs resistance.

Since our measured $S_{total}$ = 200 µV/K and $S_{oxide}$ = 400 µV/K at 300K,[30] the oxide contribution to the measured thermopower is negligible when the resistance of the oxide

$$S_{Total} = \frac{G_{oxide}S_{oxide} + G_{NW}S_{NW}}{G_{oxide} + G_{NW}}$$

is much larger than the resistance of the NWs. This can be seen by the equation below for $S_{total}$ where $G_{oxide}$ and $G_{NW}$ are the conductances of the oxide and NWs respectively. The resistance of our NWs is $\sim 10^4 \, \Omega$ at 300K.

To determine the resistance of our oxide substrate, we performed a current-voltage measurement on the substrate. Since we did not obtain a linear fit to the data (due to the tiny currents <100pA at ~ 1 V), we estimated the order of magnitude of the resistance by dividing each individual current-voltage point. This gave us a resistance order of magnitude of $\sim 10^{10} \, \Omega$ Since the substrate resistance is $\sim 10^6$ times larger than the NWs resistance and the orders of magnitude of their respective thermopowers is the same, the total thermopower, $S_{total}$, is completely dominated by the NW contribution.

## 2.5.3    Error Analysis- Thermal Conductivity

A measurement of the thermal conductivity of our Si NWs requires knowing the power dissipated in the heater and the temperature difference across the NWs. In addition, we

must determine the parasitic heat flow through the beams that support the suspended structure. Heat also flows through the thin oxide film underneath the NWs. The measured thermal conductivity value is then a contribution from the NWs + oxide. To eliminate the contribution from the oxide, we remove the NWs using a highly selective XeF$_2$ etch and determine the thermal conductivity of the oxide in a second measurement. The difference between the two measured values is the thermal conductivity of the NWs.

The thermal conductance of the sample, $K_s$, is given by (17):

$$K_s = K_b \frac{T_C}{\Delta T} \ldots\ldots\ldots\ldots \text{Equation 1}$$

where $K_b$ is the thermal conductance of the beams, $T_C$ is the temperature increase at the cold junction after the heater is turned on and $\Delta T$ is the temperature difference between the hot and cold junctions. The thermal conductance of the beams, $K_b$, is given by:

$$K_b = \frac{Q}{T_H + T_C} \ldots\ldots\ldots\ldots \text{Equation 2}$$

where $Q$ is the power dissipated in the heater and $T_H$ is the temperature increase of the hot junction after the heater is turned on. Using error propagation, the error of $K_s$, $\delta K_s$, is give by:

$$\delta K_s = K_s \sqrt{\frac{\delta K_b^2}{K_b^2} + \frac{\delta T_C^2}{T_C^2} + \frac{[\delta(T_H - T_C)]^2}{(T_H - T_C)^2}} \ldots\ldots\ldots\ldots \text{Equation 3}$$

Further simplifying this expression, we substitute for the error of $K_b$ and $\Delta T$:

$$\frac{\delta K_b^2}{K_b^2} = \frac{\delta Q^2}{Q^2} + \frac{\delta T_H^2 + \delta T_C^2}{(T_H + T_C)^2}, \ldots\ldots\ldots\ldots \text{Equation 4}$$

$$[\delta(T_H - T_C)]^2 = \delta T_H^2 + \delta T_C^2 \ldots\ldots\ldots\ldots \text{Equation 5}$$

$$\delta K_s = K_s \sqrt{\frac{\delta Q^2}{Q^2} + \frac{\delta T_C^2}{T_C^2} + 2\frac{\left(\delta T_H^2 + \delta T_C^2\right)\left(T_H^2 + T_C^2\right)}{\left(T_H^2 - T_C^2\right)^2}}$$ .............Equation 6

The error in $Q$ can be ignored because its uncertainty, $\frac{\delta Q}{Q}$, is 1 part in 1000000 or

.001%. The errors in the temperature measurement for both hot and cold junctions are

therefore the dominant terms in equation 6.  Determining the error of the temperature

measurement requires using error propagation on the formula used to find the

temperature.  The thermopower measurement section describes how we measured the

temperature of each junction.  The temperatures of the hot and cold junction are given by:

$$T_H = Q \bullet \left(\frac{dR_H/dQ}{dR_H/dT}\right) \qquad T_C = Q \bullet \left(\frac{dR_C/dQ}{dR_C/dT}\right)$$

where $R_H$ and $R_C$ are the resistances of the hot and cold thermometers.  The term,

$\left(\frac{dR/dQ}{dR/dT}\right)$ is called the conversion factor and is calculated for each thermometer.  The

terms, $dR/dQ$ and $dR/dT$, are the slopes of the best linear fits to the data points in the

resistance vs. heater power and resistance vs. temperature plots respectively (see

thermopower measurement section).

For example, each data point on the resistance vs. heater power plot is determined

by selecting a point from each plateau in the resistance vs. time plot for the heater power

calibration.  Figure 2-9 shows how this is accomplished.  Since there are 55 plateaus and

100 useable points/plateau after temperature equilibration, we generate a series of 100

lines, with each line defined by 55 points. The slopes of all 100 lines are calculated and

averaged giving $dR/dQ$. The error of the average slope is calculated by finding the

standard deviation of all 100 slopes and dividing by $\sqrt{100}$. This value is:

$$\delta\left(dR/dQ\right) = \frac{\sigma}{\sqrt{100}}$$

where $\sigma$ is the standard deviation of the 100 slope values. Similarly, $\delta\left(dR/dT\right)$ is

calculated by generating 800 linear fits, each with 40 data points (since there are 800

useable points/plateau after equilibration in the resistance vs. time plot and 40 plateaus

for the temperature calibration). The value of $dR/dT$ is found by taking the average of

all 800 slopes and the error is $\delta\left(dR/dT\right) = \frac{\sigma}{\sqrt{800}}$.

We now find the error for our temperature measurement by first calculating the

error of the conversion factor. The error in the conversion factor is:

$$\delta\left(\frac{dR/dQ}{dR/dT}\right) = \left(\frac{dR/dQ}{dR/dT}\right)\sqrt{\left[\frac{\delta\left(dR/dQ\right)}{dR/dQ}\right]^2 + \left[\frac{\delta\left(dR/dT\right)}{dR/dT}\right]^2}$$

Therefore, the error in the measured temperature for both thermometers is:

$$\delta T = T\sqrt{\left[\delta\left(\frac{dR/dQ}{dR/dT}\right)\Big/\frac{dR/dQ}{dR/dT}\right]^2}$$

where we have ignored the error in $Q$ because it is small. Based on the data collected,

our uncertainty in the temperature is .01% at 300K. This sets the uncertainty in $K_s$ to

.068%. We obtain two values for $K_s$, before etching the NWs and after etching, and then

subtract these to value to obtain $K_{NW}$. The error in $K_{NW}$ is:

$$\delta K_{NW} = \sqrt{\delta K^2_{s\,Before\,Etch} + \delta K^2_{s\,After\,Etch}}$$

We have a total error of .47 nW/K (or 9.7%) at 300K for the thermal conductance of our 10nm NWs.  Finally, we determine the thermal conductivity, $\kappa = K_s \frac{L}{A}$ where $L$ and $A$ are the length and cross-sectional area of our NWs respectively.  The uncertainty of the thermal conductivity is:

$$\delta\kappa = \kappa\sqrt{\left(\frac{\delta K_s}{K_s}\right)^2 + \left(\frac{\delta L}{L}\right)^2 + \left(\frac{\delta A}{A}\right)^2} \quad\ldots\ldots\ldots\ldots\text{equation 7}$$

The length, $L$, is measured by scanning electron microscopy (SEM) and its error is ~5nm, while $A$ is measured by AFM and its error is .1nm.  Therefore the dominant term in the right-hand side of equation 7 is the uncertainty of $K_s$.  The percent uncertainty of our thermal conductivity value at 95% confidence for our 10nm NWs is $2(\delta\kappa/\kappa)$ = 19.5% at 300K.



**Figure 2-9.  Thermal conductivity measurement data.  a.** Time-dependent plot of the resistance of the left thermometer as the heater power is ramped up 55 times.  **b.** Expanded view of data from **a**, showing three points selected in each plateau (points are exaggeratedly big).  A total of 100 points/plateau are actually selected after the resistance settled; only three are shown for clarity.  **c.** 100 linear fits to the 55 points selected.  The slopes of all 100 lines are calculated and averaged together to yield $\frac{dR}{dQ}$.  The error in $\frac{dR}{dQ}$ is found by taking the standard deviation of the slopes and dividing by $\sqrt{100}$.

## 2.5.4 Error Analysis- Thermopower and Electrical Conductivity

The measured $S$ $(= \dfrac{\Delta V}{\Delta T})$ is the slope of a best fit line to $\Delta V$ as a function of $\Delta T$. The

error of each $\Delta V$ point was found by calculating the standard deviation of each of the six

voltage measurements taken for that point (Figure 2-5A). The relative error was found to

be $\sim 1\%$ whereas for lower doped NW's, the relative error was $\sim 5\%$.

Next, the error of the temperature difference, $\delta T$, was determined using the same

method as stated above for the thermal conductivity error analysis. Since we need to

determine the temperature difference, $\Delta T$, the error is:

$$\delta T \approx \sqrt{\left[\frac{dR_H/dQ}{dR_H/dT}\cdot\sqrt{\left(\frac{\delta\left(dR_H/dQ\right)}{dR_H/dQ}\right)^2+\left(\frac{\delta\left(dR_{Ht}/dT\right)}{dR_H/dT}\right)^2}\right]^2+\left[\frac{dR_C/dQ}{dR_C/dT}\cdot\sqrt{\left(\frac{\delta\left(dR_C/dQ\right)}{dR_C/dQ}\right)^2+\left(\frac{\delta\left(dR_C/dT\right)}{dR_C/dT}\right)^2}\right]^2}$$

where we have neglected the error in $Q$ since it is less than 1%. The relative error for

$\Delta T$ is found to be 2-3% for most devices measured throughout the full temperature

range (300K to 100K).

Finally, the uncertainty in $S$ is calculated from the standard error of the slope of

the best fit lines to the $\Delta V$ vs. $\Delta T$ plot at the 95% confidence level. We find that our

relative error for $S$ is between 1and 6%.

The error of the electrical conductivity is less than 1% and the current-voltage

traces are highly linear with $R^2 \sim .999$.

## 2.5.5　Error Analysis– ZT

The uncertainty in $ZT$ is calculated using the following formula:

$$\delta ZT = ZT \sqrt{\left[ 4\left(\frac{\delta S}{S}\right)^2 + \left(\frac{\delta \kappa}{\kappa}\right)^2 + \left(\frac{\delta \sigma}{\sigma}\right)^2 \right]}.$$

The factor of four appears in the above equation because $ZT \propto S^2$.  The relative error in $\sigma$

can be ignored since it is less than 1%.  The uncertainty in $ZT$ at the 95% confidence

level, $2 \times \delta ZT$.

## 2.6　References

1. Y. M. Lin, X. Z. Sun, M. S. Dresselhaus, *Phys. Rev. B* **2000**, *62*, 4610.
2. Y. Liu, R. E. Allen, *Phys. Rev. B* **1995**, *52*, 1566.
3. X. Sun, Z. Zhang, M. S. Dresselhaus, *Appl. Phys. Lett.* **1999**, *74*, 4005.
4. Chen, G. *et al*. Recent developments in thermoelectric materials. *Int. Mater. Rev.* **48**, 45-66 (2003)
5. Dresselhaus, M.S. *Physics of the Solid State* **41**, 679 (1999).
6. Heremans, J.P., Thrush, C.M., Morelli, D.T. & Wu, M.C. *Physical Review Letters* **88**, 216801 (2002).
7. Venkatasubramanian, R. *et al*. Thin-film thermoelectric devices with high room-temperature figures of merit. *Nature* **413**, 597-602 (2001).
8. Zhang, Z.B., Ying, J.Y. & Dresselhaus, M.S. *Journal of Materials Research* **13**, 1745 (1998).
9. Heremans, J. *et al*. *Physical Review B* **58**, R10091 (1998)
10. Liu, K.. *et al*. *Applied Physics Letters* **73**, 1436 (1998).
11. Zhang, Z.B. *et al*. *Applied Physics Letters* **73**, 1589 (1998).
12. Heremans, J. *et al*. *Physical Review B* **61**, 2921 (2000).
13. Lin, Y.M. *et al*. *Applied Physics Letters* **76**, 3944 (2000).
14. Hong, K.M. *et al*. *Journal of Applied Physics* **85**, 6184 (1999).
15. Lin, Y.M. *et al*. *Applied Physics Letters* **81**, 2403 (2002).
16. Heremans, J. & Thrush, C.M. *Physical Review B* **59**, 12579 (1999).
17. Dresselhaus, M.S. *et al*. *Semiconductors and Semimetals* **71**, 1 (2001).
18. Cronin, S.B. et al. in *Materials Research Society Symposium Proceedigs Volume 582* (Eds.: S. T. Pantelides, M. A. Reed, J. S. Murday, A. Aviram), Materials Research Society Press, Pittsburgh, H10.4.1 (2000).
19. Small, J.P., Perez, K.M. & Kim, P. *Physical Review Letters* **91**, 256801 (2003).
20. Llaguno, MC., Fischer, J.E., Johnson, A.T. & Hone, J. *Nano Letters* **4**, 45 (2004).

21. J. Chang, H. Kim, J. Han, M. H. Jeon, W. Y. Lee, *J. Appl. Phys.* **2005**, *98*, 023906.
22. Wang, D., Sheriff, B.A. & Heath, J.R. Complementary symmetry silicon nanowire logic: Power-efficient inverters with gain. *Small* **2**, 1153-1158 (2006).
23. Melosh, N.A. *et al*. Ultra-high density nanowire lattices and circuits. *Science* **300**, 112-115 (2003).
24. Llaguno, M.C., Fischer, J.E., Johnson A.T. & Hone, J. Observation of thermopower oscillations in the coulomb blockade regime in a semiconducting carbon nanotube. *Nano Lett.* **4**, 45-49 (2004).
25. Zhou, F. *et. al.*, Determination of transport properties in chromium disilicide nanowires via combined thermoelectric and structural characterizations. *Nano Letters* **7**, 1649-1654 (2007).
26. Li, S *et al*. Measuring thermal and thermoelectric properties of one-dimensional nanostructures using a microfabricated device. *Journal of Heat Transfer* **125**, 881-888 (2003).
27. Li, D. *et al*. Thermal conductivity of individual silicon nanowires. *Applied Physics Letters* **83**, 2934-2936 (2003).
28. Williams, K.R. & Muller, R.S. Etch rates for micromachining processing. *Journal of Microelectromechanical Systems* **5**, 256-269 (1996).
29. Liu, W.J. & Asheghi, M. Thermal conduction in ultrahigh pure and doped single-crystal silicon layers at high temperatures. *Journal of Applied Physics* **98**, 123523 (2005).
30. Beynon, J. & Steele, C.B. Thermopower measurements on $SiO_x$ thin films. Journal of Materials Science 25, 4255-4258 (1990).

## Appendix 2.1    Thermopower Measurement

A thermoelectric measurement can be setup in such a way that there are only two regions

that have temperature gradients (Figure 2-10).  Applying Kirchoff's voltage law to the

circuit in Figure 2-10 yields the following:



Figure 2-10
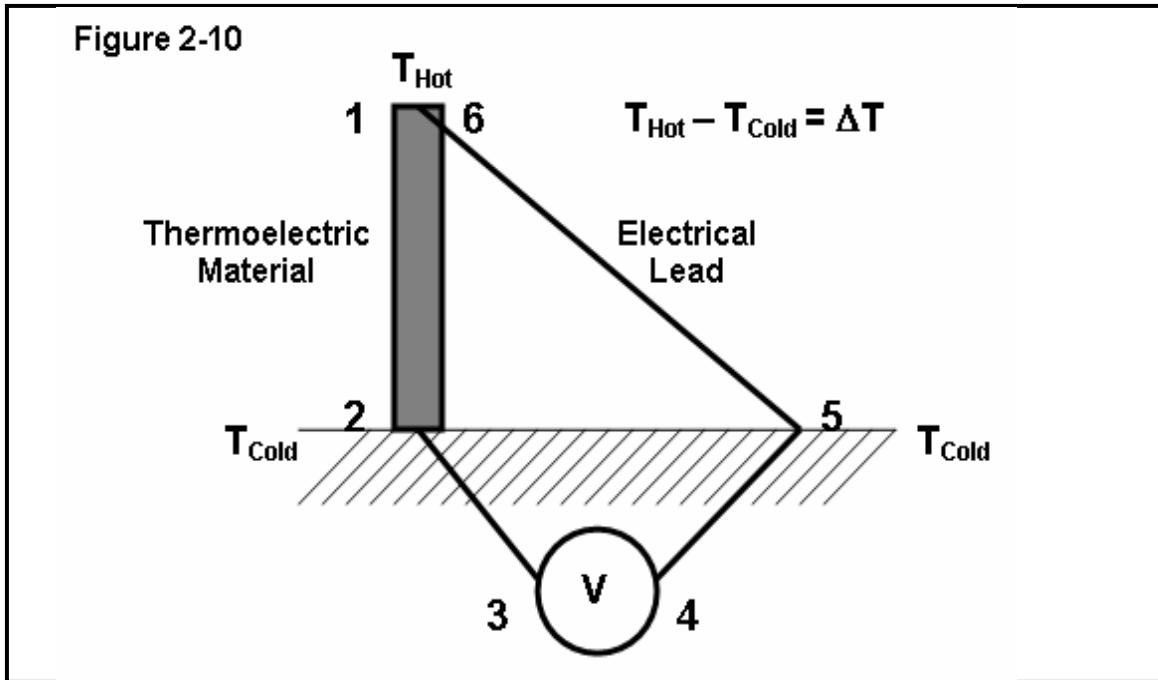
1. $V = \int_1^2 E \cdot dx + \int_2^3 E \cdot dx + \int_4^5 E \cdot dx + \int_5^6 E \cdot dx$

Substituting for the electric field $E = S\nabla T$

2. $V = \int_1^2 S \frac{dT}{dx} \cdot dx + \int_2^3 S \frac{dT}{dx} \cdot dx + \int_4^5 S \frac{dT}{dx} \cdot dx + \int_5^6 S \frac{dT}{dx} \cdot dx$

3. Simplifying: $V = \int_1^2 SdT + \int_2^3 SdT + \int_4^5 SdT + \int_5^6 SdT$

4. The only temperature gradients exist between points 1 and 2 and 5 and 6.  Therefore:

$$V = S_{thermoelectric}\Delta T - S_{lead}\Delta T$$

# Chapter 3

# Thermoelectric Results on Bismuth Nanowires

## 3.1    4–Point Resistance Results and Discussion

Resistivity of the NWs can be calculated from the four-point resistance since the NW dimensions and numbers of NWs measured are both known. The resistivity at 300K is 13, 9.5, 55, and $66\,\mu\Omega\cdot m$ for cross sections of 15 nm by 72, 55, 40 and 28 nm, respectively, which corresponds to 10, 7, 42, and 50 times of that of bulk Bi. The increase of resistivity for thinner Bi wires reflects greater contributions from boundary and surface scatterings. Previously reported four-point measurement on a 70 nm diameter NW prepared from template assisted growth yielded a resistivity 6 times greater than that of bulk Bi[1].

The temperature dependent resistance is given in Figure 3-1(a). The observed trends generally agree with two-point measurements on vapor-deposited Bi nanowire arrays:[2] The widest wires yield bulk-like semimetal behavior, while the resistance for the thinnest two sets of NWs increases monotonically with decreasing temperature. The 55 nm wires exhibit an intermediate behavior. These trends have been cited as evidence that a quantum confinement-induced semimetal-to-semiconductor transition occurs when the wire width decreases below about 50 nm,[3,4] consistent with theoretical predictions[5]. However, separating changes in carrier density and carrier mobility is inherently

**Figure 3-1. Electrical transport and thermoelectric properties of Bi nanowires. (a).** The temperature depen-dent resistance of Bi NWs, normalized to the resistance at 300K. Each point in the plot is obtained by fitting a four-point *I-V* trace. **(b).** Absolute values of measured thermoelectric voltage of Bi NWs at 300K, as a function of temperature difference across the nanowires. Thermoelectric power **S** is obtained from the slope of a linear fit to the data. (Note: 55nm NWs yielded a positive thermoelectric power while others gave a negative thermoelectric power.) **(c).** Thermoelectric power **S** as a function of temperature for different NWs.

difficult[6]. Our electric field effect measurements (see below) suggest the 40 nm wide NWs are still semimetallic. Consequently, the observed trends may reflect a decreased mobility through surface and grain-boundary scattering[3].

A resistance drop is observed below 10K in all devices. This is reminiscent of the reported superconducting transition of amorphous Bi films around 6K,[7,8] and may imply the existence of amorphous regions within the polycrystalline structure. We found that the drop disappears when a strong magnetic field is applied, which presumably quenches the superconducting transition.

## 3.2　Thermopower Results and Discussion

The thermopower measurements on the Bi NWs are shown in Figure 3-1(b) and 3-1(c). Figure 3-1(b) demonstrates the linearity of the thermoelectric voltage measured with respect to the applied temperature difference, indicating $S = \dfrac{\mathrm{d}V}{\mathrm{d}T}$ is a constant. The measured $S$ is really that of the NW/electrode couple since it contains a contribution from the thermoelectric power of the metal electrodes, $S_m$. However, the $S$ of Pt and Ti tend to cancel each other and are both within $\pm 5\,\mu V/K$ at all temperatures studied here[9].

The 72 nm wide NWs show the largest (room temperature) $S$ of $-25.2\ \mu V/K$. This value is about one half of that of bulk Bi in the same direction (perpendicular to the trigonal axis)[10], but agrees with both earlier measurements of $S$ ($-28\ \mu V/K$) of thin (33 nm) Bi film[10] and recent results ($-22\ \mu V/K$ and $-28\ \mu V/K$) on two NW arrays[11]. $S$ is reduced dramatically for the smaller NWs. Both theory[12] and experiment[10] have concluded that surface and grain-boundary scattering will reduce $S$ in polycrystalline thin films. This may also apply to the width for the NWs. The magnitude of $S$ for all NWs tends to zero as the temperature is decreased. $S$ is also the entropy per electric charge, and so must go to zero at 0 K.

For bulk Bi, and for the 72, 40, and 28 nm NWs $S$ is negative valued, because electrons have a much higher mobility than holes and dominate the transport[13,14]. The 55 nm NW, on the other hand, shows a positive $S$　[Fig. 3-1(c)] indicating p-type carrier dominated transport, which has been reported in two Bi NW composites and a $200 \pm 50\,nm$ Bi wire[15] before. Calculations based on the gating response of our NWs (see

below) suggest that an impurity level of 10 ppm could be sufficient to change $S$ from negative to positive.

## 3.3   Gating Response of Bismuth Nanowires

Controlling the position of the Fermi energy ($E_F$) in order to enhance the thermoelectric power[4] may be achieved through doping or the electric field effect (EFE)[16,17]. Controlling the doping level to a precision of $10^{17}\,\text{cm}^{-3}$ (or 4 ppm of dopants) is desirable, but hard to achieve experimentally[5]. For example, 99.999% purity Bi has an impurity level of 10 ppm. However, our Bi NWs are thinner than the screening length,[17] thus enabling EFE measurements. The effective doping level can be estimated based on the Si-SiO$_2$-Bi capacitance structure as a function of the gating voltage $V_g$:

$$N_d = \frac{Q_d}{eV_{Bi}} = \frac{CV_g}{eAh_{Bi}} = \frac{\varepsilon_0\varepsilon_{r,SiO_2}\dfrac{A}{d_{SiO_2}}}{eAh_{Bi}}V_g = \frac{\varepsilon_0\varepsilon_{r,SiO_2}}{ed_{SiO_2}h_{Bi}}V_g \quad (1)$$

where $\varepsilon_0\varepsilon_{r,SiO_2}$ is the permittivity of the SiO$_2$ layer, which serves as the dielectric medium, and $d_{SiO_2} = 500\,\text{nm}$ and $h_{Bi} = 15\,\text{nm}$ are the thickness of the SiO$_2$ layer and the height of the Bi NWs, respectively. From Eq.(1) we obtain $N_d = 2.8\times10^{16}\text{cm}^{-3}/\text{V}\cdot V_g$. Thus, the effective dopant concentration can be controlled to $10^{16}\,\text{cm}^{-3}$ or lower with appropriate $V_g$. Our devices functioned properly up to $V_g = \pm150\text{V}$, or an effective dopant concentration of $4\times10^{18}\,\text{cm}^{-3}$. This is sufficient for the thermoelectric optimization predicted by theory[5]. The EFE results of the 40 nm Bi NW device at 20 K are given in Fig 3-2. Ohmic conductivity was observed for all gating voltages, and the

conductance (*G*) values were obtained from a linear fit to the *I-V* curves. Similar behavior has been observed in thin carbon films[17].

The conductance and thermoelectric power trends in Fig 3-2, which were collected separately during several different scans to avoid possible drift, are found to be well correlated by the Mott formula, which relates the diffusive thermoelectric power, $S_d$, to the conductance $G$[18]:

$$S_d = \frac{-\pi^2 k_B^2 T}{3|e|} \frac{1}{G} \frac{dG}{dE}\bigg|_{E=E_F} \quad (2)$$

An increase in $V_g$ would put more electrons into the Bi nanowires and elevate $E_F$, *i.e.*

$\dfrac{dE_F}{dV_g}$ is positive. Consequently, $\dfrac{dG}{dV_g} = \dfrac{dE_F}{dV_g} \dfrac{dG}{dE}\bigg|_{E=E_F}$ is expected to be of the same sign

as $\dfrac{dG}{dE}\bigg|_{E=E_F}$, and thus from Eq. (2) $S_d$ should have the opposite sign of $\dfrac{dG}{dV_g}$. The

measured *S* clearly shows this trend: a sudden drop in the measured *S* is observed when the slope of *G* with respect to $V_g$ changes from negative to positive at point A. This correlates to a change of positive $S_d$ to negative $S_d$. Another decrease in *S* (and an increase in the slope of *G*) is observed around point B. It naturally follows that somewhere around point A (~ −10 V), with a measured *S* of between −0.5 to −0.4 μV/K, should correspond to an $S_d$ value of 0. The observed non-zero value of *S* is reflective of two facts: 1) $S_g$, the phonon-drag component of *S* has not been taken into account, which could be significant at this temperature[19], and 2) $S_m$, the thermoelectric power of Pt/Ti electrodes used to establish the thermocouple with Bi nanowires has not been considered.

The Pt/Ti electrodes are not expected to contribute to the EFE due to their extremely small screening lengths[17], but they might cause a uniform shift $-S_\mathbf{m}$ in $S$ at all $V_\mathbf{g}$.
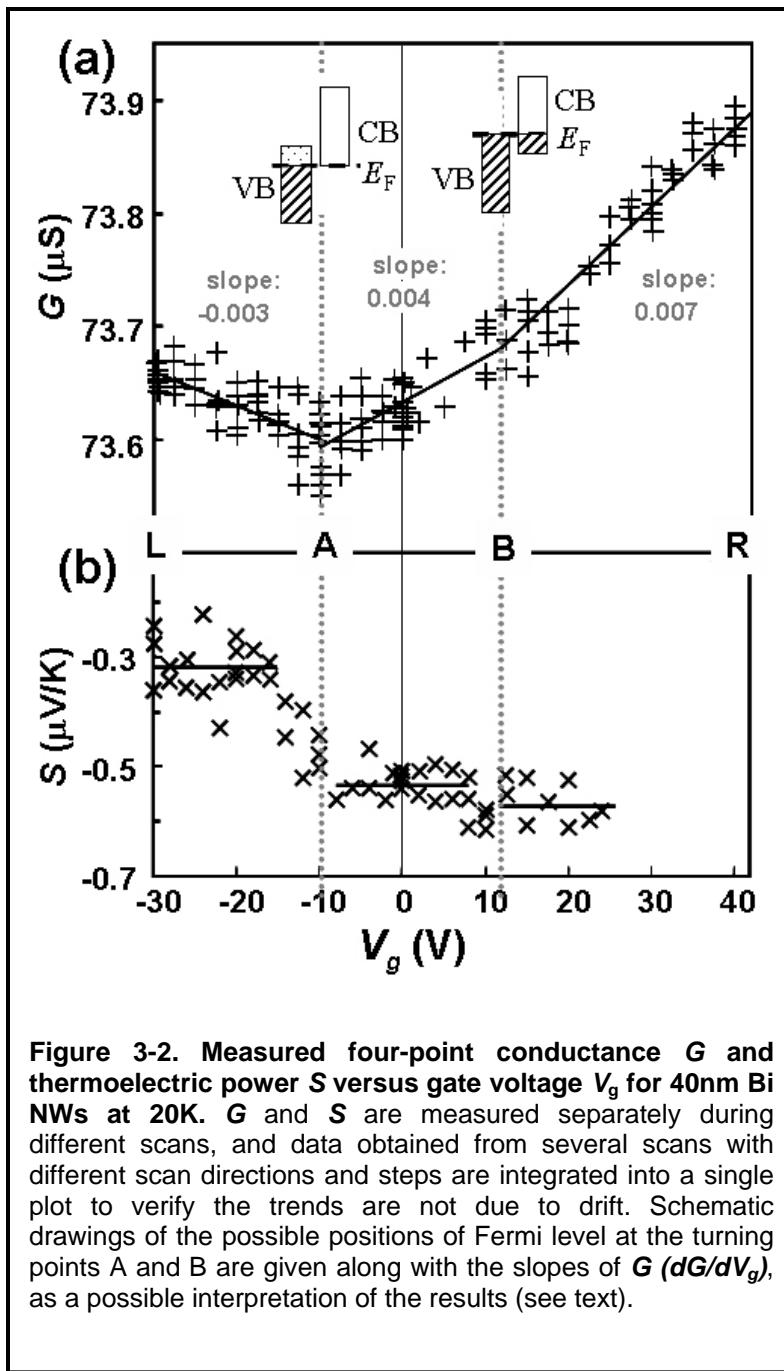


**Figure 3-2. Measured four-point conductance $G$ and thermoelectric power $S$ versus gate voltage $V_g$ for 40nm Bi NWs at 20K. $G$ and $S$ are measured separately during** different scans, and data obtained from several scans with different scan directions and steps are integrated into a single plot to verify the trends are not due to drift. Schematic drawings of the possible positions of Fermi level at the turning points A and B are given along with the slopes of $G$ $(dG/dV_g)$, as a possible interpretation of the results (see text).

40 nm Bi NWs have been predicted to be semiconductors at 20 K for any growth direction[5], and results similar to Figure 3-1(a) has been cited as evidence of the semimetal-to-semiconductor transition around 55 nm[3,20]. Our EFE result, however, is found to be more readily explained with a semimetal model[17]. The two turning points A and B around ±10 V correspond to an effective dopant concentration of about

$N_d = 3 \times 10^{17}\,\mathbf{cm^{-3}}$, which is close to the carrier concentrations measured in bulk (semimetal) Bi, $N_e \approx N_h = 2.5 \sim 4 \times 10^{17}\,\mathbf{cm^{-3}}$, obtained from various methods[21,22].

These two turning points are thus assigned to be when the Fermi level is at the bottom of the conduction band (A) and at the top of the valence band (B). If the bottom of conduction band is lower than the valence band as in a semimetal [see drawings in Fig 3-2(a)], the observed results can be rationalized. According to the assignment, holes are the only carrier when $V_g < -10$ V, which results in positive $S_d$, and a decrease in $G$ with an increase in $V_g$, as the concentration of holes would decrease when putting more electrons in the nanowires. Similarly, electrons are the only carrier for $V_g > 12$ V, which results in negative $S_d$ and an increase in $G$ with increase of $V_g$. For $-12$ V $< V_g < 12$ V, both electrons and holes carriers are present and contribute to the conductivity $\sigma$ and thermoelectric power $S$:

$$\sigma = \sigma_e + \sigma_h = e(N_e \mu_e + N_h \mu_h) \qquad (3)$$

$$S = \frac{\sigma_e S_e + \sigma_h S_h}{\sigma_e + \sigma_h} = \frac{N_e \mu_e S_e + N_h \mu_h S_h}{N_e \mu_e + N_h \mu_h} \qquad (4)$$

An increase in $V_g$ in this region would cause an increase in $N_e$, the concentration of electrons, but a decrease in $N_h$, the concentration of holes.

Consider the value of $dG/dV_g$ [the slope of Fig. 3-2(a)] in the three regions L-A, A-B, and B-R. Note that $|slope|_{B-R} > |slope|_{L-A}$. This is consistent with the higher mobility[13,14] electrons dominating the region B-R and holes dominating L-A. Also note that $slope_{A-B} \approx slope_{L-A} + slope_{B-R}$. This indicates that both electrons and holes participate in the charge transport in the region A-B. A more quantitative analysis of the results would require knowledge of the density of states (DOS), but the near-linear dependence of $G$ over $V_g$ in the different regions suggests a near-constant DOS characteristic of a 2D

system[17]. This is reasonable for a $15\,\text{nm} \times 40\,\text{nm}$ cross section wire as the electronic de Broglie wavelength is 30~40 nm in Bi[23].

Our EFE results suggest semimetal behavior for 40nm Bi NWs, with electron and hole carrier concentrations comparable to bulk values, contradicting theoretical predictions that Bi nanowires thinner than 50 nm would become semiconductor at this temperature[5], and that as a consequence the total carrier density for 40 nm NW at 20 K would decrease significantly to smaller than $10^{16}\,\text{cm}^{-3}$, about 100 times smaller than the bulk value[5]. This could stem from the polycrystalline nature of our NWs, since interface states would be expected to contribute to the total density of states near $E_\text{F}$. Similar conclusions have been drawn from recent Shubnikov-de Haas oscillation measurements on 30 nm Bi wires in terms of surface-induced charge carriers[6]. Our 28 nm NW device was found to give similar trends as the 40 nm NW device, but the data set was less complete.

## 3.4   Conclusion

In conclusion, nanofabrication methods and a device architecture have been combined to allow for four-point electric, thermoelectric, and electric field effect measurements on individual Bi nanowires. No clear semimetal-to-semiconductor transition or enhancement in thermoelectric power has been observed, probably due to the polycrystalline nature of our NWs and the presence of surface states. Therefore, the passivation of such surface states and single-crystalline material are necessary to improve the thermoelectric performance of Bi nanowires.

## 3.5    References

1.  S. B. Cronin, Y. M. Lin, T. Koga, J. Y. Ying, M. S. Dresselhaus, in Mat. Res. Soc. Symp. Proc., Vol. 582 (Eds.: S. T. Pantelides, M. A. Reed, J. S. Murday, A. Aviram), Materials Research Society Press, Pittsburgh, p. H10.4.1 (2000).
2.  J. Heremans, C. M. Thrush, Y. M. Lin, S. B. Cronin, Z. Zhang, M. S. Dresselhaus, J. F. Mansfield, *Phys. Rev. B* **2000**, *61*, 2921.
3.  Y. M. Lin, S. B. Cronin, J. Y. Ying, M. S. Dresselhaus, J. P. Heremans, *Appl. Phys. Lett.* **2000**, *76*, 3944.
4.  M. S. Dresselhaus, Y. M. Lin, S. B. Cronin, O. Rabin, M. R. Black, G. Dresselhaus, T. Koga, *Semiconduct. Semimet.* **2001**, *71*, 1.
5.  Y. M. Lin, X. Z. Sun, M. S. Dresselhaus, *Phys. Rev. B* **2000**, *62*, 4610.
6.  T. E. Huber, A. Nikolaeva, D. Gitsu, L. Konopko, C. A. Foss, M. J. Graf, *Appl. Phys. Lett.* **2004**, *84*, 1326.
7.  W. Buckel, R. Hilsch, *Z. Physik* **1954**, *138*, 109.
8.  J. S. Shier, D. M. Ginsberg, *Phys. Rev.* **1966**, *147*, 384.
9.  C. L. Foiles, in *Landolt-Börnstein - Group III Condensed Matter, Vol. 15, Subvol. B* (Eds.: K. H. Hellwege, J. L. Olsen), Springer-Verlag, Berlin, **1985**, pp. 48.
10. V. Damodara Das, N. Soundararajan, *Phys. Rev. B* **1987**, *35*, 5990.
11. J. Heremans, C. M. Thrush, *Phys. Rev. B* **1999**, *59*, 12579.
12. C. R. Pichard, C. R. Tellier, A. J. Tosser, *J. Phys. F: Met. Phys* **1980**, *10*, 2009.
13. D. L. Partin, J. Heremans, D. T. Morelli, C. M. Thrush, C. H. Olk, T. A. Perry, *Phys. Rev. B* **1988**, *38*, 3818.
14. A. H. Dekuijper, J. Bisschop, *Thin Solid Films* **1983**, *110*, 99.
15. A. D. Grozav, E. Condrea, *J. Phys.: Condes. Matter* **2004**, *16*, 6507.
16. V. Sandomirsky, A. V. Butenko, R. Levin, Y. Schlesinger, *J. Appl. Phys.* **2001**, *90*, 2370.
17. K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, A. A. Firsov, *Science* **2004**, *306*, 666.
18. M. Cutler, N. F. Mott, *Phys. Rev.* **1969**, *181*, 1336.
19. M. W. Wu, N. J. M. Horing, H. L. Cui, *Phys. Rev. B* **1996**, *54*, 5438.
20. Y. M. Lin, O. Rabin, S. B. Cronin, J. Y. Ying, M. S. Dresselhaus, *Appl. Phys. Lett.* **2002**, *81*, 2403.
21. J. P. Michenau, J. P. Issi, *J. Phys. C: Solid State Phys.* **1972**, *5*, 3061.
22. G. A. Williams, *Phys. Rev.* **1965**, *139*, A771.
23. N. Garcia, M. Strongin, Y. H. Kao, *Phys. Rev. B* **1972**, *5*, 2029.

# Chapter 4

# Thermoelectric Results on Silicon Nanowires[1]

## 4.1    Introduction

The most efficient TEs have historically been heavily doped semiconductors because the Pauli principle restricts the heat carrying electrons to within $kT$ of the Fermi level[2] for metals.  The Wiedemann-Franz law, $(\kappa_e/\sigma T) = (\pi^2/3)(k/e)^2 = (156 \ \mu V/K)^2$, where $\kappa_e$ is the electronic contribution to $\kappa$, constrains $ZT$.  Semiconductors have a lower density of carriers, leading to larger $S$ and $\kappa$ that is dominated by phonons ($\kappa_{ph}$) implying that the electrical and thermal conductivities are somewhat decoupled[2].   $\kappa$ can be reduced by using bulk semiconductors of high atomic weight, which decreases the speed of sound. However this strategy has not yet produced materials with $ZT > 1$.

For a metal or highly doped semiconductor, $S$ is proportional to the energy derivative of the density of electronic states (DOS).  In low-dimensional (nanostructured) systems the DOS has sharp peaks[6-8] and, theoretically, a high thermopower. Harnessing this electronic effect to produce high $ZT$ materials has only had limited success[9,10]. However, optimization of the phonon dynamics and heat transport physics in nanostructured systems has yielded results[3-5]. Nanostructures may be prepared with one or more dimensions smaller than the mean free path of the phonons and yet larger than that of electrons and holes.  This potentially reduces $\kappa$ without decreasing $\sigma$[11].  Bulk Si is

a poor thermoelectric $(ZT_{300K} \sim 0.01)^{23}$, and this phonon physics is important for our SiNWs where the electronic structure remains bulk-like.
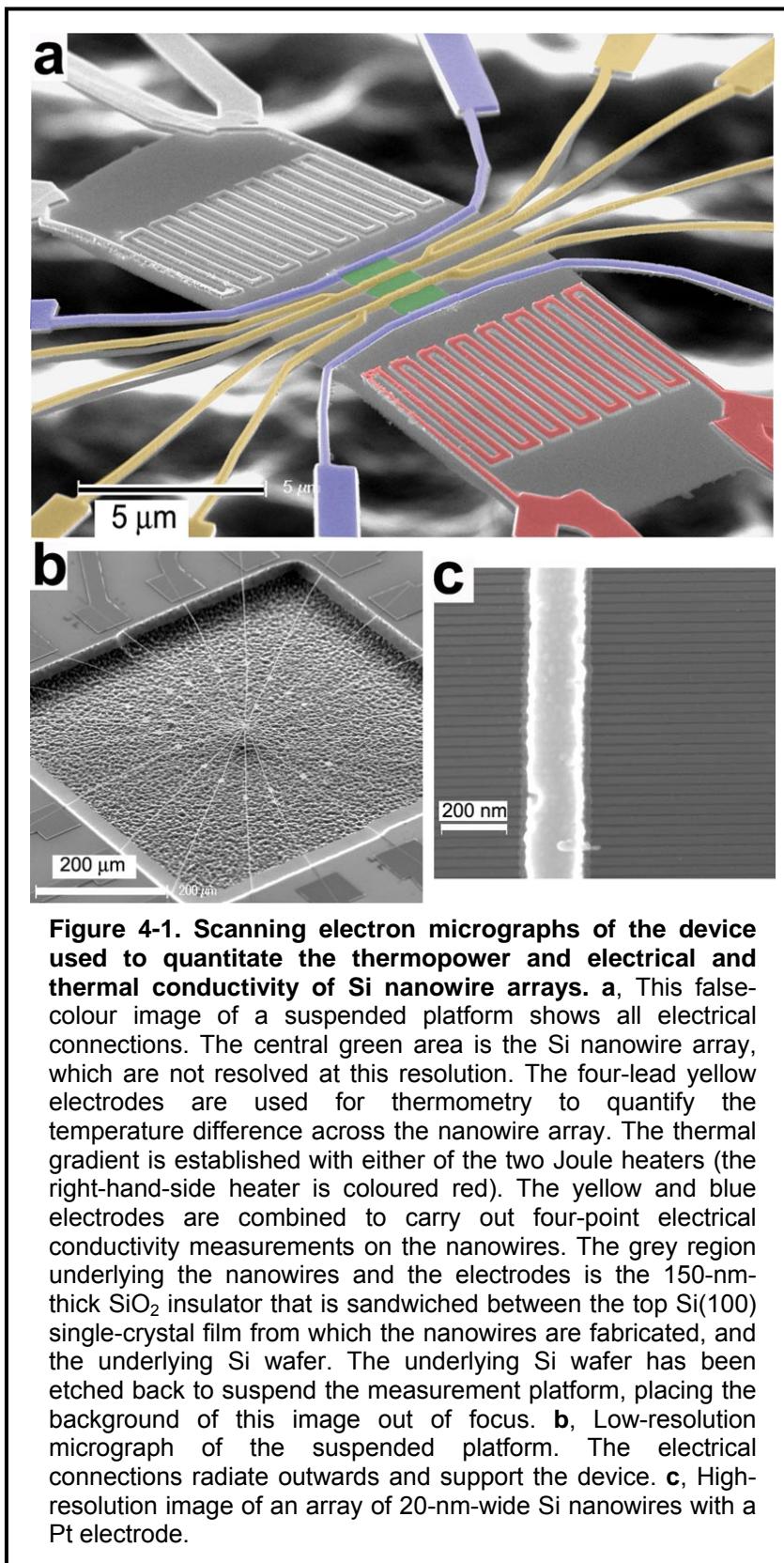


**Figure 4-1. Scanning electron micrographs of the device used to quantitate the thermopower and electrical and thermal conductivity of Si nanowire arrays. a**, This false-colour image of a suspended platform shows all electrical connections. The central green area is the Si nanowire array, which are not resolved at this resolution. The four-lead yellow electrodes are used for thermometry to quantify the temperature difference across the nanowire array. The thermal gradient is established with either of the two Joule heaters (the right-hand-side heater is coloured red). The yellow and blue electrodes are combined to carry out four-point electrical conductivity measurements on the nanowires. The grey region underlying the nanowires and the electrodes is the 150-nm-thick $SiO_2$ insulator that is sandwiched between the top Si(100) single-crystal film from which the nanowires are fabricated, and the underlying Si wafer. The underlying Si wafer has been etched back to suspend the measurement platform, placing the background of this image out of focus. **b**, Low-resolution micrograph of the suspended platform. The electrical connections radiate outwards and support the device. **c**, High-resolution image of an array of 20-nm-wide Si nanowires with a Pt electrode.

## 4.2 Results of S, $\sigma$, and $\kappa$

Figure 4-1 shows images of the devices and the Si NWs we utilized for these experiments. The platform permits 4-point measurements of the electrical conductivity of the NWs, Joule heating to establish a thermal gradient across the NWs, and 4-point thermometry to quantify that gradient[9,12]. The resistance of the 4-point thermometry

electrodes is typically 2 orders of magnitude smaller than the resistance of the NWs. The measurement platform is suspended in vacuum to allow measurement of NW thermal conductivity[13,14]. For all measurements, the SiNWs could be selectively removed using a XeF$_2$ etch, thus allowing for measurements of the contributions to $\sigma$, $S$, and $\kappa$ from the platform and oxide substrate.

All values of $S$, $\sigma$, and $\kappa$ reported here are normalized to individual NWs, although each experiment utilized a known number of NWs ranging from 10 and 400. The Si microwires and NWs were prepared using the same substrates, doping methods, etc., but different patterning methods (electron-beam lithography versus SNAP). Measurements of $\sigma$, $\kappa$ and $S^2$ for Si NWs (and microwires) for different NW sizes and/or doping levels are presented in Figures 4-2 and 4-3. The NW electrical conductivity is between 10 and 90% of the bulk, depending on NW dimensions. A reduced $\sigma$ likely arises from surface scattering of charge carriers[15]. Nevertheless, all NWs are highly doped and most exhibit metallic-like conductivity (increasing $\sigma$ with decreasing $T$).
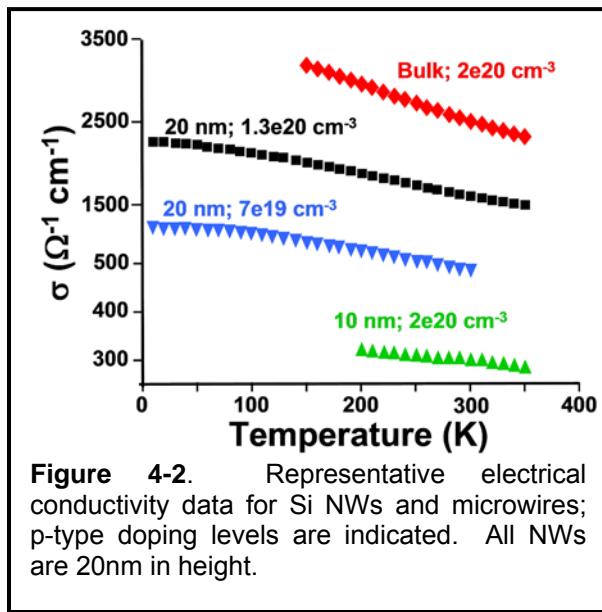


**Figure 4-2.** Representative electrical conductivity data for Si NWs and microwires; p-type doping levels are indicated. All NWs are 20nm in height.

The temperature dependence of $\kappa$ for a microwire and 10 and 20 nm wide NWs were recorded at modest statistical resolution in order to establish trends. This data indicated that $\kappa$ drops sharply with shrinking NW cross section (Figure 4-3a) and that the 10 nm wide NWs exhibited a $\kappa$

$(0.76 \ W \cdot m^{-1} \cdot K^{-1} \ +/-0.15 \ W \cdot m^{-1} \cdot K^{-1})$ that was below the theoretical limit of $0.99 \ W \cdot m^{-1} \cdot K^{-1}$



**Figure 4-3 Factors contributing to *ZT* for various Si nanowires. All nanowires are 20 nm in height. a**, The temperature dependence of the thermal conductivity $\kappa$, presented as $\kappa_{bulk}/\kappa_{nanowire}$ to highlight the improvement that the reduction of $\kappa$ in nanowires lends to ***ZT***. $\kappa_{bulk}$ values, which are slightly below the true bulk value for Si, are taken from an identically measured 520 nm × 35 nm-sized film. The inset scanning electron microscope micrographs show the region of the device containing the nanowires before (top) and after (bottom) the $XeF_2$ etch to remove the nanowires. **b**, The temperature dependence of $S^2$ for 20-nm-wide Si nanowires at various p-type doping concentrations (indicated on the graph). Note that the most highly doped nanowires (pink line) yield a thermopower similar to that of bulk Si doped at a lower level. For nanowires doped at slightly higher and slightly lower concentrations than the bulk, *S* peaks near 200 K. This is a consequence of the one-dimensional nature of the Si nanowires.

for bulk Si. Thus, very large data sets were collected for 10 and 20 nm wide NWs to allow for a more precise determination of $\kappa$. Our observed high ***ZT*** for Si NWs (Figure 4-4) occurs because $\kappa$ is sharply reduced and the phonon drag component of the thermopower, $S_{ph}$, becomes large. Below, we show that $S_{ph}$ increases due to a 3D to 1D dimensional crossover of the phonons participating in phonon drag, and to decreasing $\kappa$. However, we first address the issue that our measured $\kappa$ at 300K for 10 nm wide Si NWs is $< \kappa_{min}$[16].

**Figure 4-4 Temperature dependence of *ZT* for two different groups of nanowires.** The cross-sectional area of the nanowires, and the p-type doping level, are given. The 20-nm-wide nanowires have a thermopower that is dominated by phonon contributions, and a *ZT* value ~1 is achieved near 200 K. The smaller (10-nm-wide) nanowires have a thermopower that is dominated by electronic contributions. The *ZT* at 350 K is calculated using the thermal conductivity value for the 10-nm-wide nanowires at 300 K. The error bars represent 95% confidence limits.

## 4.2.1  Discussion of $\kappa_{min}$

The derivation of $\kappa_{min}$ assumes that the minimum path length of wavelength $\lambda$ phonons is $(1/2)\lambda$ and that the phonons are described by the Debye model using bulk sound speeds with no optical modes. The $(1/2)\lambda$ value is an order of magnitude estimate and is difficult to determine precisely in analogy to the minimum electron mean free path used to calculate the Mott-Ioffe-Regel $\sigma_{min}$. Also, $\kappa_{min}$ is proportional to the transverse and longitudinal acoustic speeds of sound[16]. These are reduced in our NWs at long wavelengths because the modes become 1D. The ratio of the 1D to 3D longitudinal speeds of sound is $[(1+v)(1-2v)/(1-v)]^{1/2} = 0.87$ where $v = 0.29$ is the Poisson ratio of Si. The transverse acoustic speed goes to zero at long wavelength since $\omega \sim k^2 d$ where $d$ is the NW width[17]. Thus the bulk $\kappa_{min}$ estimate above is invalid for our NWs and values smaller than $\kappa_{min}$ are attainable.

## 4.2.2   Phonon Drag

For all but the most highly doped NWs, $S$ peaks near 200K (Figure 4-3b). This peak is unexpected: similarly doped bulk Si exhibits a gradual decrease in $S$ as $T$ is reduced (green trace). For $T < 100K$ a peaked $S(T)$ is observed for metals and lightly doped semiconductors and is due to phonon drag[18-20].

Phonon drag is generally assumed to vanish with decreasing sample dimensions because the phonon path-length is limited by the sample size[21-23]. This appears to eliminate phonon drag as the reason for the peak in our NWs. We show below that the phonon wavelengths participating in drag are on the order of or larger than the wire width. This leads to a 3D to 1D dimensional crossover of these modes and removes the cross-sectional wire dimensions from limiting the phonon mean path (see Figure 4-5 inset). The NW boundaries are incorporated into the 1D mode and are not an obstacle to phonon propagation. Therefore, the limiting size becomes the wire length ($\sim 1$ $\mu$m) and phonon drag "reappears" at very small dimensions. In addition, classical elasticity theory[17] is valid for the phonon wavelengths considered here[24], leading to thermoelastic damping[17,25,26] of sound waves proportional to $\kappa$. Thus $S_{ph}$ is further enhanced due to the observed reduced thermal conductivity $\kappa$.

It might appear that elasticity theory leads to a contradiction because $\kappa$ is proportional to the mean phonon lifetime. If the phonon lifetimes increase as stated above, then $\kappa$ should also increase. The resolution is that the elasticity expression is only valid for long wavelength modes. We now consider separately the electronic and phonon contributions, $S=S_e+S_{ph}$ to the thermopower for the NW data at $T > 200K$.
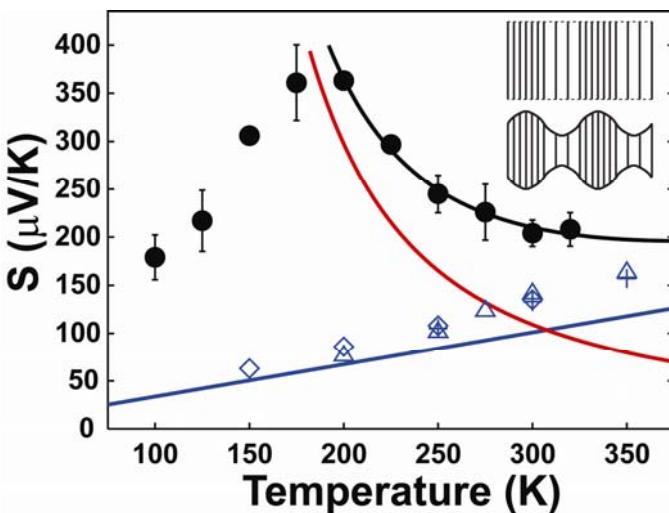
**Figure 4-5. Thermopower calculation plotted along with experimental data (black points) from a 20-nm-wide Si nanowire p-type doped at 3 × 10$^{19}$ cm$^{-3}$.** The black curve is the fitted expression for the total thermopower $S_e + S_{ph}$. The red curve is the phonon contribution $S_{ph}$ and the blue line is the electronic term $S_e$ arising from the fit. The fit has maximum error 6.1 μV K$^{-1}$ and root-mean-square error 1.8 μV K$^{-1}$. The experimental error bars represent 95% confidence limits and at 150, 200 and 225 K are smaller than the data points. The blue data points are experimental values for bulk wires (doping 2 × 10$^{20}$ cm$^{-3}$; crosses), 10-nm-wide nanowires (doping 7 × 10$^{19}$ cm$^{-3}$; diamonds), and 20-nm-wide wires (doping 1.3 × 10$^{20}$ cm$^{-3}$; triangles) where only a linear-$T$ electronic contribution was found. This data are close to the extracted electronic contribution from the black data points (blue line) and shows that the fitted linear term is reasonable. The drop in **S** to 0 as $T{\to}0$ occurs because the phonon mean free path reaches the sample size and the specific heat tends to 0 according to the third law of thermodynamics. The inset shows the character of a three-dimensional bulk longitudinal acoustic phonon mode (top) and a one-dimensional mode when the wavelength is larger or of the order of the width (bottom). The one-dimensional mode incorporates the existence of the boundary by transverse expansion (or compression) for longitudinal compression (or expansion). The ratio of the transverse strain to the longitudinal strain is the Poisson ratio (0.29 for Si).

Charge carriers dissipate heat to the lattice through a process that first involves momentum conserving (non-dissipative) electron-phonon collisions. The phonons that contribute to phonon drag cannot have a wavelength shorter than $\lambda_{min}$, which is determined by the size of the Fermi surface. Phonon drag is observed in metals only at low $T$ because the Fermi surface is large and the heat carrying short wavelength phonons have short lifetimes. At low $T$ (<20K), $S_{ph} \sim T^3$ from the phonon specific heat ($\sim T^3$). For $kT \gg \Theta_{Debye}$, the specific heat becomes

constant and the number of phonons available for phonon-phonon scattering is $\sim T$ leading to $S_{ph} \sim 1/T$ (1).

For p-type Si, the holes are near the valence band maximum. The phonon drag modes are acoustic with largest wavevector, $k_{ph} = 2k_{fermi} = 0.2$ Å$^{-1}$ (for impurity doping 3 x $10^{19}$ cm$^{-3}$). The shortest wavelength is $\lambda_{ph} = 2\pi/k = 31$ Å. Umklapp (non-momentum conserving) phonon-phonon scattering processes determine the rate of phonon heat dissipation. The Debye energy ($\Theta_D$) sets the energy scale for Umklapp scattering. The number of Umklapp phonons available to dissipate the long wavelength phonons is given by the Bose-Einstein function

$$N_U = \frac{1}{e^{\Theta_D/T} - 1}$$

leading to a scattering rate $1/\tau_{ph} \sim N_U$. When $T \gg \Theta_D$, $1/\tau_{ph} \sim T$. Since $\Theta_{Debye} = 640$K for Si, the full Bose-Einstein expression must be applied for $T \leq 350$K.

The electronic contribution ($S_e$) is estimated from the Mott formula (1)

$$S_e(T) = \frac{\pi^2 k^2 T}{3e} \left( \frac{\partial \ln \sigma(\varepsilon)}{\partial \varepsilon} \right) \approx (283 \, \mu V \, / \, K)(kT \, / \, E_f),$$

where the conductivity derivative equals the reciprocal of the energy scale over which it varies (the Fermi energy $E_f$). Assuming hole doping occurs in the heavier Si valence band (mass 0.49), this leads to $E_f = 0.072$ eV $= 833$K and $k_f = 0.1$ Å$^{-1}$ for $n = 3$ x $10^{19}$ cm$^{-3}$. Thus $S_e(T) = aT$ where $a = 0.34$ µV/K$^2$.

The $T > 200$K thermopower data of the 20 nm wire (doping n=3x$10^{19}$cm$^{-3}$) fits

$$S = S_e + S_{ph} = aT + b\left[\exp\left(\Theta_D / T\right) - 1\right]$$

where **a**, **b** and **$\Theta_D$** are varied to obtain the best fit (Figure 4-5). The coefficients are **a** = 0.337 $\mu$V/K$^2$, **b** = 22.1 $\mu$V/K, and **$\Theta_{Debye}$** = 534 K. The coefficient **a** is almost identical to our estimate of 0.34 $\mu$V/K$^2$. Thus phonon drag explains the observed thermopower. Consistent with measurements of phonon drag in bulk Si[21], **S** in our NWs increases significantly at lighter doping. This data, plus a fit of the Figure 4-5 data with **$\Theta_{Debye}$** fixed to experiment at 640K, is presented in section 4.3.4.

The phonon drag contribution to **S** is of the form[20,27]

$$S_{ph} \sim \left(\frac{\tau_{ph}}{\mu T}\right).$$

**$\tau_{ph}$**, the phonon lifetime, is ~ 1/**$\kappa$** from elasticity theory. $\mu$ is the electron mobility. **ZT** scales as (neglecting **$S_e$**)

$$S_{ph} \sim \frac{1}{\mu T \kappa}, \quad \sigma \sim n\mu, \quad ZT \sim \frac{n}{\mu T \kappa^3},$$

leading to increased **ZT** with *decreasing* mobility. This is opposite the conclusion reached when considering only **$S_e$**[6].

## 4.3    Theoretical Analysis

### 4.3.1    1D versus 3D electronic structure.

The first question to consider in understanding the unexpectedly large Seebeck effect observed for the relatively lightly doped NWs is whether the electronic structure in the 10-

20 nm wires is 1D or 3D above 100K. This can be estimated from the magnitude of the zero-point energy due to the finite cross-section of the wire. The effective masses for the two bands at the Valence Band Maximum of Si are 0.49 and 0.16 electron masses. Hole doping will occur predominantly in the heavier band. The zero-point energy is

$$\varepsilon = \frac{\hbar^2 k^2}{2m^*}, \quad k = \frac{\pi}{d},$$

where d is the wire diameter. For $m^* = 0.49$, $\varepsilon = 89$K for $d = 10$ nm and $\varepsilon = 22$K for $d = 20$ nm. Thus for temperatures above 100K thermal smearing is larger than the zero-point energies, suggesting that the electronic properties should be 3D. Furthermore, the thermopower data in Figure 4-5 (green and pink data points) and 4-3 (blue data points) shows that the 10 and 20nm wires and bulk samples with similar impurity doping levels have an almost identical magnitude and linear temperature dependence. Since these 10 nm, 20 nm, and bulk wires have similar thermopowers, the different zero-point energies are not leading to a large change in $S$ and we may conclude the electronic structure in all the wires remains 3D for temperatures > 100K.

## 4.3.2    Is there a Boron impurity band?

For Boron doping levels within the range of $10^{19}$-$10^{20}$ cm$^{-3}$, the electron gas becomes degenerate and it is possible that a B impurity band forms containing the charge carriers. By comparing an estimate of the thermopower due to holes in the Si valence bands of the NW with measured values, we conclude that this is unlikely.

If the holes are primarily in the Si valence bands, then the holes are predominantly in the heavier valence band with mass 0.49. For doping $n$, the Fermi momentum satisfies

$$n = \frac{k_f^3}{3\pi^2}.$$

For $n = 3 \times 10^{19}$ cm$^{-3}$, $k_f = 0.1$ Å$^{-1}$ leading to Fermi energy $E_f = 0.072$ eV $= 833$K. From the Mott formula for the electronic contribution $S$:

$$S_e(T) = \frac{\pi^2 k^2 T}{3e}\left(\frac{\partial \ln \sigma(\varepsilon)}{\partial \varepsilon}\right) \approx (283 \, \mu V / K)(kT / E_f),$$

we estimate that $S_e = 102$ μV/K at $T = 300$K. We assumed that the derivative of the logarithm of the conductivity is the reciprocal of the energy scale over which $\sigma(\varepsilon)$ changes $\sim 1/E_f$ in computing $S_e$. This estimate for $S_e$ is compatible with $S_e = 130$-$170$ μV/K from experiments for bulk samples at 300K for wires that only have a linear T electronic contribution to $S$. A Boron impurity band would likely lead to much heavier masses and hence to $S_e$ larger than observed.

## 4.3.3    Derivation of the results from elasticity theory for the phonon lifetime

For doping $n = 3 \times 10^{19}$ cm$^{-3}$, the longitudinal acoustic (LA) phonons that can receive momentum from the electrons are characterized by wavelengths of $> 31$ Å, corresponding to $> 5.5$ unit cells of Si. Classical elasticity theory models the system by an elastic continuum. Indeed, recent molecular dynamics simulations show classical elasticity is a good approximation for these wavelengths[24]. Since the wavelengths of the phonons

involved in phonon drag are larger than several unit cells, classical elasticity theory is expected to give a good approximation to their dynamics. When a sound wave propagates through an elastic medium it generates local temperature gradients from the local volume changes in the wave. Thermal conduction acts to remove these gradients by heat flow with a corresponding increase in entropy. This irreversible heat flow leads to damping of the sound wave[23].

Consider a volume element $\delta x \delta y \delta z = \delta x\, \delta A$ with a temperature gradient along the x-direction. The temperature at $x$ is $T$ and $T + \delta T$ at $x + \delta x$. Over a time interval $\delta t$ an amount of heat $\delta Q$

$$\delta Q = (+\kappa \nabla T)\delta A\, \delta t$$

flows to $T$. $\kappa$ is the thermal conductivity. The change in entropy ($S$) is

$$\delta S = \frac{\delta Q}{T} - \frac{\delta Q}{T + \delta T}.$$

The irreversible energy lost due to the increasing entropy is

$$T\delta S = \frac{\delta T}{T}\delta Q,$$

leading to the rate of energy lost by thermal conduction

$$\frac{dE}{dt} = -\frac{\delta T}{T}\frac{\delta Q}{\delta t} = -\frac{\delta T}{T}\kappa \frac{\delta T}{\delta x}\delta y \delta z = -\frac{\kappa}{T}\left(\frac{\delta T}{\delta x}\right)^2 dV,$$

$$\frac{dE}{dT} = -\kappa \int \frac{1}{T}(\nabla T)^2\, dV.$$

A 3D longitudinal wave given by $u_x = u_0 cos(kx-\omega t)$ where $\omega = c_L k$ and $c_L$ is the longitudinal speed of sound leads to local temperature gradients due to local volume changes as the wave propagates. The volume deformations of the longitudinal wave are separated by one-half the wavelength $\lambda$. For low $\omega$, the separation between the high and low temperature regions is too large to be equalized during the period of the wave. Thus for low $\omega$, the volume changes are adiabatic[28]. At high frequencies, the wavelength is small and thermal gradients are rapidly equalized leading to isothermal volume changes for large $\omega$[28].

A 3D longitudinal acoustic (LA) phonon with momentum ($q_x$, $q_y$, $q_z$) that makes an angle $\Theta$ with the long axis of the 1D wire (z-direction)  is a linear combination of a longitudinal phonon of momentum $q_z$ with amplitude cos$\Theta$ and a transverse acoustic (TA) phonon with momentum $q_z$ with amplitude sin$\Theta$.  Thus electron-phonon scattering around the Fermi surface leads to a 1D LA mode with probability $\cos^2\Theta$ and a TA mode with probability $\sin^2\Theta$. TA modes have temperature gradients perpendicular to the wire axis. LA modes have temperature gradients along the wire axis.

The largest momentum (shortest wavelength) phonon modes participating in phonon drag by being excited due to the holes scattering with phonons are 3D longitudinal acoustic with wavevector, $k_{LA} = 2k_f = 0.2$ Å$^{-1}$ for an impurity doping level of 3 x 10$^{19}$ cm$^{-3}$. The wavelength is $\lambda_{sh} = 2\pi/k = 31$ Å. The speed of sound in the (100) direction for Si is $c_L = 8.43 \times 10^5$ cm/s, leading to $\omega_{LA} = c_L k = 1.7$ x 10$^{13}$ sec$^{-1}$. Phonon drag thermopower is an average over all the phonon modes that are scattered by the electrons at the Fermi surface. Thus $\omega_{LA}$ is an upper bound on the phonon frequencies participating in phonon drag.

The crossover frequency from adiabatic volume deformations to isothermal deformations for LA modes occurs at $\omega_0(LA) = c_L^2/\chi$ where $\chi$ is the thermal diffusivity. $\chi = 0.8$ cm$^2$/s for bulk Si at 300K leading to $\omega_0(LA) = 8.9$ x $10^{11}$ sec$^{-1}$. 3D LA modes with frequency $\omega_{LA}$ are isothermal.

For TA modes, the crossover frequency is determined by the width of the wire, $d$, and the thermal diffusivity, $\omega_0(TA) = \pi^2\chi/d^2$. In this case, frequencies *below* $\omega_0(TA)$ are isothermal and frequencies *above* $\omega_0(TA)$ are adiabatic[25,26]. For $d = 20$nm, $\omega_0(TA) = 2$ x $10^{12}$ s$^{-1}$ using the bulk value of the thermal diffusivity.

For the NW, the thermal diffusivity is ~100 times smaller than bulk since the diffusivity is proportional to the thermal conductivity. Thus the LA crossover frequency is ~100 times larger than the bulk $\omega_0(LA)$ leading to adiabatic damping for LA modes. The TA mode crossover frequency decreases by a factor of ~ 100. TA modes are also adiabatic in the NW at 300K.

The damping rate for an LA sound wave of momentum $k$ is[28,29]

$$\frac{1}{\tau_{LA}(k)} = \frac{1}{\tau_0} \cdot \frac{(\omega_k\tau_0)^2}{1+(\omega_k\tau_0)^2} \cdot \varepsilon,$$

where the terms in the expression are defined as

$$\omega_k = c_L k, \quad \tau_0 = \frac{\chi}{c_L^2}, \quad \chi = \frac{\kappa}{\rho C_p}, \quad \varepsilon = \frac{v_{th}^2}{c_L^2}, \quad v_{th}^2 = \left[\frac{\alpha^2 K_{ad}^2 T}{\rho C_p}\right].$$

$\rho$ is the density, $\chi$ is the thermal diffusivity, $\kappa$ is the thermal conductivity, $\alpha$ is the volume thermal expansion coefficient ($\alpha = 3\alpha_{linear}$ where $\alpha_{linear}$ is the linear expansion

coefficient) , $K_{ad}$ is the adiabatic bulk modulus, $c_L$ is the speed of sound, and $C_p$ is the constant pressure specific heat per volume. For $\omega_k \gg 1/\tau_0$, the damping saturates to $\varepsilon/\tau_0$. This occurs for 3D LA modes relevant to phonon drag.

The ratio of the 1D and 3D LA scattering times for a mode with momentum $k$ is

$$\frac{\tau_{3D}(k)}{\tau_{1D}(k)} = (kc_L)^2 \left(\frac{\chi}{c_L^2}\right)^2 \left[\frac{(1-v)}{(1+v)(1-2v)}\right]\left[\frac{\kappa(1D)}{\kappa(3D)}\right],$$

where $v$ is the Poisson ratio ($v = 0.29$ for Si). For the largest momentum $k = 2k_f$, the 3D LA mode lifetime is 3.3 times larger than the 1D LA mode. We take the ratio of the thermal conductivities to be 1/100. For the shortest wavelength LA modes, phonon drag is not enhanced in the wire. For long wavelength LA modes (small $k$), the lifetime of the 1D mode rapidly increases relative to the bulk due to the $k^2$ term and phonon drag is enhanced. In addition, the 3D scattering rate for long wavelength LA modes becomes adiabatic rather than isothermal leading to a further increase in the 1D LA mode lifetime relative to 3D. In this regime, the lifetime is increased by the ratio of the 3D to 1D thermal conductivities or $\sim 100$.

The scattering rate of the 1D TA modes is smaller than the corresponding LA modes in bulk. The scattering rate of the 1D TA mode is[28]

$$\frac{1}{\tau_{TA}(k)} = \omega_0(TA) \cdot \frac{(\omega_k/\omega_0(TA))^2}{1+(\omega_k/\omega_0(TA))^2} \cdot \Delta, \qquad \Delta = \left[\frac{\alpha_{linear}^2 K_{ad} T}{C_p}\right]$$

Since $\omega \gg \omega_0(TA)$, the TA modes are adiabatic and the scattering rate saturates to $\Delta\omega_0$. Using experimental values $K_{ad} = 9.8 \times 10^{11}$ dyne/cm, $\alpha_{linear} = 2.6 \times 10^{-6}$ K$^{-1}$, and $\rho = 2.3$ g/cm$^3$ leads to $\tau_{3D}/\tau_{TA} \sim 1/2$. The phonon drag contribution from the TA modes is

increased due to the decreased thermal conductivity in the NW. Thus we have shown that the decreased thermal conductivity leads to increased phonon drag due to TA modes and long wavelength LA modes.

### 4.3.4    Theoretical Fit to the Phonon Drag Data

In Figure 4-6 we show how well the high temperature data (>200K) in Figure 4-3b is explained by:

$$S = S_e + S_{ph} = aT + b\left[\exp\left(\Theta_D / T\right) - 1\right]$$



where we set the Debye energy parameter to its experimental value of $\Theta_{Debye}$ = 640K. As shown in Figures 4-5 and 4-6, phonon drag explains the observed monotonic increase in *S* as *T* decreases.
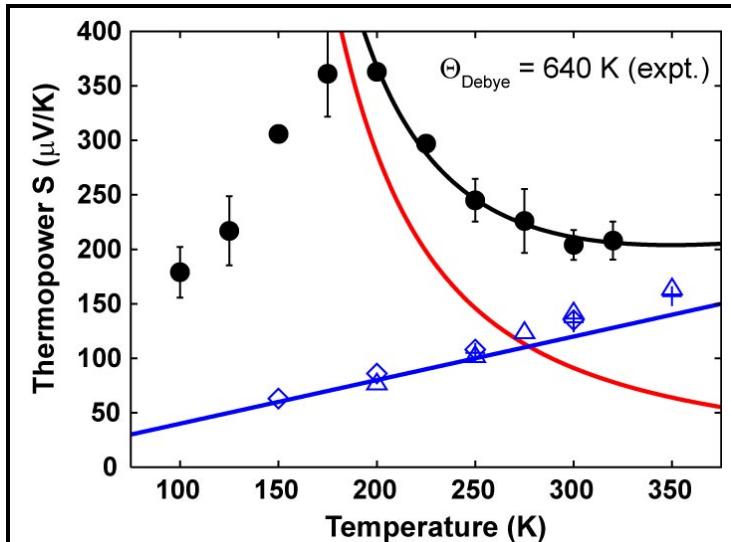
**Figure 4-6. A refit of Fig. 4 using the experimentally determined $\Theta_{Debye}$ = 640K.** The points are the experimental data also presented in Fig. 4. The black, red and blue lines represent the fitted curves of *S*, *S$_{ph}$*, and *S$_e$*, respectively, utilizing the expression in the text. We find, *a* = 0.40 µV/K$^2$ and *b* = 12.2 µV/K. The fitted maximum and rms error are 8.9 µV/K and 2.1 µV/K respectively. The experimental error bars at 150, 200, and 225K are smaller than the data points. The fitted electronic coefficient value is close to our estimated value of *a* = 0.34 µV/K$^2$ assuming hole doping of Si with no boron impurity band. It is also close to the blue symbols in this figure which are for samples that only exhibit a linear *T*, *S$_e$* component.

Ultimately, we must have **S**→0 as T →0 (by the 3$^{rd}$ law of thermodynamics) and this is observed experimentally. One factor

that we have not included is that at sufficiently low temperature, the mean free path approaches the dimensions of the system leading to scattering that eventually reduces $S$ toward zero. It is also possible that for the doping levels used here the energy dependence of the electron-phonon scattering reduces the momentum transfer to LA phonons leading to a further reduction in $S_{ph}$ at low $T$.

## 4.4    Conclusion

We have combined experiment and theory to demonstrate that semiconductor NWs can be designed to achieve extremely large enhancements in thermoelectric efficiency, and we have shown that the temperature of maximum efficiency may be tuned by changing the doping and the NW size. Theory indicates that similar improvements should be achievable for other semiconductor NW systems because of phonon effects. These nanowire thermoelectrics may find applications related to on-chip heat recovery, cooling, and power generation. Additional improvements through further optimization of NW size, doping, composition, etc., should be possible.

## 4.5    References

1.  This text of this chapter is taken from:  Boukai, A.I. *et al*. Silicon nanowires as efficient thermoelectric materials.  *Nature* (accepted 11/02/07).

2.  MacDonald, D.K.C. *Thermoelectricity: An Introduction to the Principles* (Wiley, New York, 1962).

3.  Venkatasubramanian, R. *et al*. Thin-film thermoelectric devices with high room-temperature figures of merit. *Nature* **413**, 597-602 (2001).

4.  Harman, T.C. *et al*. Quantum dot superlattice thermoelectric materials and devices. *Science* **297**, 2229-2232 (2002).

5.  Hsu, K.F. *et al*. Cubic AgPb$_m$SbTe$_{2+m}$: Bulk thermoelectric materials with high figure of merit. *Science* **303**, 818-821 (2004).
6.  Hicks, L.D. & Dresselhaus, M.S. Thermoelectric figure of merit of a one-dimensional conductor. *Physical Review B* **47**, 16631-16634 (1993).
7.  Mahan, G.D. & Sofo, J.O. The best thermoelectric. *PNAS* **93**, 7436-7439 (1996).
8.  Humphrey, T.E. & Linke, H. Reversible thermoelectric nanomaterials. *Physical Review Letters* **94**, 096601 (2005).
9.  Boukai, A., Xu., K. & Heath, J.R. Size-dependent transport and thermoelectric properties of individual polycrystalline bismuth nanowires. *Advanced Materials* **18**, 864-869 (2006).
10. Yu-Ming, L. *et al*. Semimetal-semiconductor transition in Bi$_{1-x}$Sb$_x$ alloy nanowires and their thermoelectric properties. *Applied Physics Letters* **81**, 2403-2405 (2002).
11. Majumdar, A. Enhanced: Thermoelectricity in semiconductor nanostructures. *Science* **303**, 777-778 (2004).
12. Small, J.P., Perez, K.M. & Kim, P. Modulation of thermoelectric power of individual carbon nanotubes. *Physical Review Letters* **91**, 256801 (2003).
13. Li, S *et al*. Measuring thermal and thermoelectric properties of one-dimensional nanostructures using a microfabricated device. *Journal of Heat Transfer* **125**, 881-888 (2003).
14. Li, D. *et al*. Thermal conductivity of individual silicon nanowires. *Applied Physics Letters* **83**, 2934-2936 (2003)
15. Wang, D., Sheriff, B.A. & Heath, J.R. Complementary symmetry silicon nanowire logic: Power-efficient inverters with gain. *Small* **2**, 1153-1158 (2006).
16. Cahill, D. G., Watson, S.K., & Pohl, R.O. Lower limit to the thermal conductivity of disordered crystals. *Physical Review B* **46**, 6131-6140 (1992).
17. Landau, L.D. & Lifshitz, E.M. in *Theory of Elasticity, 3$^{rd}$ Edition,* 138 (Butterworth Heinemann, Oxford, 1986).
18. Pearson, W.B. Survey of thermoelectric studies of the group-1 metals at low temperatures carried out at the national-research-laboratories, ottawa. *Soviet Physics-Solid State* **3**, 1024-1033 (1961).
19. Weber, L. & Gmelin, E. Transport properties of silicon. *Applied Physics A: Solids and Surfaces* **53**, 136-140 (1991).
20. Herring, C. Theory of the thermoelectric power of semiconductors. *Physical Review* **96**, 1163-1187 (1954).
21. Geballe, T.H. & Hull, G.W. Seebeck Effect in Silicon. *Physical Review* **98**, 940-947 (1955).
22. Behnen, E. Quantitative examination of the thermoelectric power of n-type Si in the phonon drag regime. *Journal of Applied Physics* **67**, 287-292 (1990).
23. Trzcinksi, R., Gmelin, E. & Queisser, H.J. Quenched Phonon Drag in Silicon Microcontacts. *Phys. Rev. Lett.* **56**, 1086-1089 (1986).
24. Maranganti, R. & Sharma, P. Length scales at which classical elasticity breaks down for various materials. *Physical Review Letters* **98**, 195504 (2007).
25. Lifshitz, R. & Roukes, M.L. Thermoelastic damping in micro- and nanomechanical systems. *Physical Review B* **61**, 5600-5609 (2000).

26. Zener, C. Internal friction in solids I. Theory of internal friction in reeds. *Physical Review* **52**, 230-235 (1937).

27. Gurevich, L. Thermoelectric properties of conductors *J. Phys.* (U.S.S.R) **9**, 477 (1945).

28. Deresiewicz, H. Plane waves in a thermoelastic solid. *Journal of the Acoustical Society of America* **29**, 204-209 (1957).

29. Chadwick, P. & Sneddon, I.N. Plane waves in an elastic solid conducting heat. *Journal of the Mechanics and Physics of Solids* **6**, 223-230 (1958).

# Chapter 5

# Future Directions

## 5.1    Reversible Thermoelectrics

The figure of merit, **ZT**, must be large to obtain a high efficiency.  Principally, the thermopower and electrical conductivity must be large, and the thermal conductivity must be small.  This general approach requires testing many different types of binary and even ternary materials.    Until recently, there has been no known materials independent approach to reach the Carnot limit.  A paper, published by T.E. Humphrey and H. Linke explains how to engineer a nanostructured material that would potentially provide an infinite **ZT** and thus, a route towards the Carnot efficiency[1].

The Carnot limit of any heat engine is achieved only when the heat engine operates under reversible equilibrium conditions[2].  For a particle heat engine, such as a thermoelectric device where the working fluid consists of charged carriers, the equilibrium condition is satisfied when the Fermi-Dirac function is spatially invariant across the entire thermoelectric material[1].    The Fermi function, $f = \dfrac{1}{e^{(E-\mu/kT)}+1}$,

describes the probability that an electronic state is occupied.  If the temperature, **T**, and chemical potential, $\mu$, are constant across the material, the $f$ will be invariant.  There is no preferred direction for particle flow or energy and therefore the electronic system is under equilibrium.    Under such reversible equilibrium conditions, no additional entropy,

stemming from frictional or resistive losses is produced and the heat engine operates at the Carnot limit.  The challenge, therefore, is to engineer a thermoelectric material in a way that provides for the spatial invariance of $f$ under the imposed conditions of a temperature and chemical potential difference across that material[3].

A thermoelectric device operates under the conditions of a temperature and chemical potential difference.  In addition, bulk materials possess a continuum of electronic states with energies, $E$, that the charged carriers are allowed to occupy.  Under a temperature gradient, the charged carriers will occupy a different distribution of those states that is appropriate at each temperature along the gradient.  Therefore, $f$ in a thermoelectric material is not spatially invariant.  However, a nanostructured material composed of a superlattice nanowire (SLNW) has the ability to allow $f$ to be constant across the material even under the influence of temperature and chemical potential differences.

A SLNW is a material composed of a series of quantum dots connected together,
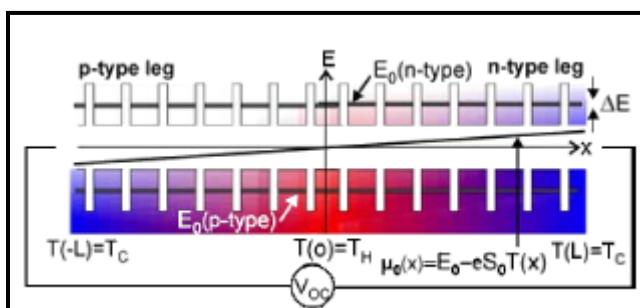


**Figure 5-1. A superlattice nanowire (SLNW) consists of a series of quantum dots each with its own miniband $\Delta E$.** The entire material is under equilibrium when $\Delta E$ approaches a single energy state and the chemical potential and temperature differences offset each other exactly.  These conditions make f spatially invariant. Figure is from Reference 1.

each with its own miniband $\Delta E$ for electrons or holes.  $\Delta E$ represents a continuum of electronic states that the electrons or holes occupy within the quantum well.  The chemical potential is tuned, either by doping or a gate voltage, across the SLNW to counteract the temperature difference so that the term, ($E$ –

$\mu$)/$T$, is constant across the entire length of the SLNW when $E$ represents only one electric energy level (Figure 5-1). If the quantum dots are made small so that the electronic states are separated by an energy larger than $kT$, then the electrons or holes occupy a single energy level $\Delta E = E_0$. The density of states is essentially composed of delta functions separated by an energy larger than $kT$ so that only one state is occupied for each quantum dot in the SLNW. These conditions place the thermoelectric material under equilibrium and allow the efficiency to approach the Carnot limit.

## 5.1.1    Silicon Nanowire Based Quantum Dot Superlattices

A SLNW can be fabricated from a Si NW using the SNAP process. First, SNAP is used to create an array of degenerately doped Si NWs 20 nm in width and 20 nm tall. Next, a closely spaced array of Pt NWs fabricated by SNAP with a pitch of 10 nm and NW width of 20 nm is placed atop the Si NWs. Next, the entire structure is subjected to an $O_2$ plasma using a reactive ion etcher operating at 100W and 20 sccm for 5 minutes. Finally the Pt NWs are removed by placing the structure in a boiling solution of aqua regia for 20 minutes. The result is an array of Si NWs that have been periodically oxidized every 20 nm. This effectively creates a series of Si quantum dots 20 **x** 20 **x** 20 nm separated by a 10 nm thick tunneling barrier composed of $SiO_2$.

Atomic force microscopy (AFM) was used to characterize the Si quantum dot superlattice. Specifically, an AFM was operated in non-contact mode using electric force microscopy (EFM) whereby a +5V voltage is applied to a conductive Si tip. The tip hovers 50 nm above the surface of the sample as it is scanned. The grounded sample deflects the tip by either a repulsive or attractive force which depends on the sample's

conductance.  When the tip is over the conductive Si quantum dots, there is an attractive force between it and the sample.  When the tip is over the insulating $SiO_2$ tunneling barrier, the attractive force is diminished.  This can be seen in the EFM image in Figure 5-2.



**Figure 5-2.  AFM measurement on Si/SiO$_2$ superlattice nanowires.**  A)  EFM measurement on the Si/SiO$_2$ superlattice nanowires with the tip held at +5V.  The peaks represent the Si regions while the valleys represent the SiO$_2$ regions.  B)  AFM topography measurement on the same structure in A).  The vertical scale bar represents 2 nm indicating that the surface is nearly flat.
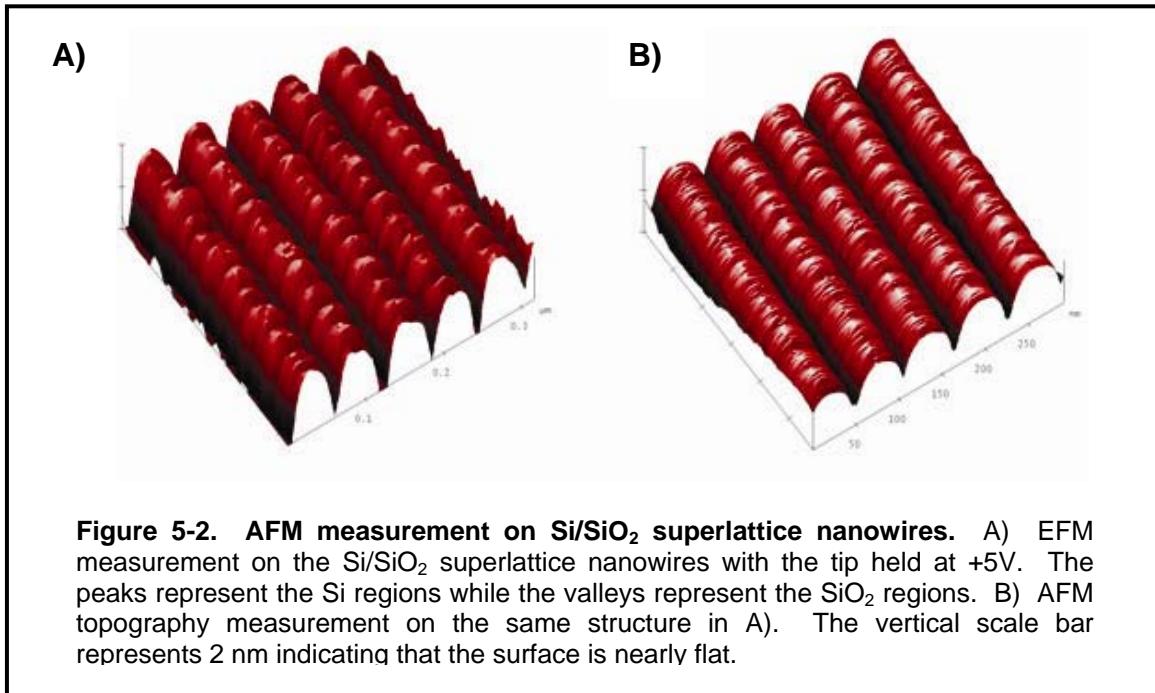
The image in 5-2A is an EFM scan with the tip held at +5V.  The peaks are Si and the valleys are the oxidized Si surfaces.  The image in Figure 5-2B was taken in tapping mode and shows the topography of the surface.  The vertical scale bar in Figure 5-2B is 2 nm which clearly indicates that the surface is effectively flat.  Therefore, the peaks in Figure 5-2A originate from the conductance difference of the Si and oxide and are not artifacts from topographic Van der Waals forces.  Figure 5-3 shows a two dimensional EFM image along with the periodicity of the Si quantum dots.  The periodicity of the Si dots closely follows the periodicity of the original SNAP lattice that was used to fabricate

the device, which was ~30 nm. The device was not tested electrically due to time constraints.



**Figure 5-3. A 2-D EFM image of the superlattice nanowire along with its spectral period.** The spectral period closely matches that of the original SNAP superlattice used to make the structure.
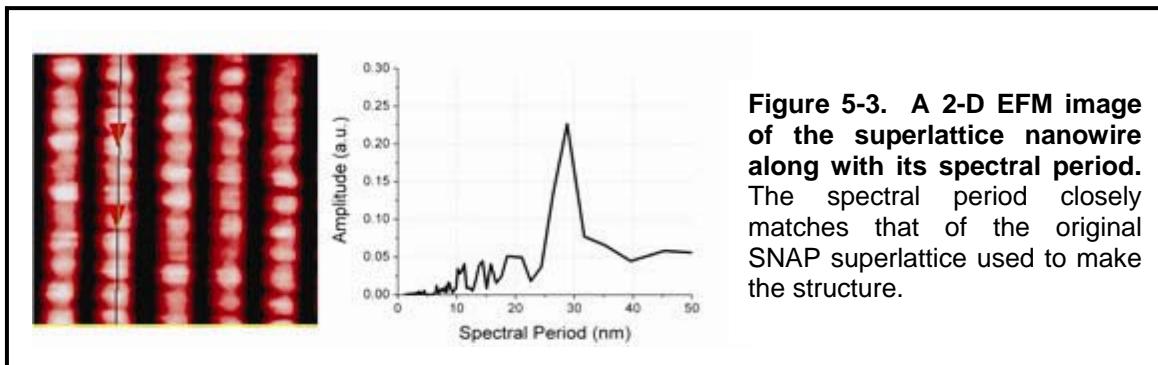
## 5.2 Enhanced Thermoelectric Properties of Thin Silicon Films With Nanostructured Holes

Bulk Si would exhibit a large figure of merit if its thermal conductivity was not so high. One strategy to increase ZT would be to alloy Si with germanium to decrease the thermal conductivity. This has already been tried and Si/Ge alloys find use as thermoelectrics at high temperatures where the thermopower is higher[4]. Another strategy to make bulk Si a useful thermoelectric material at room temperature would be to introduce defects that scatter phonons yet do not deleteriously affect the electrical transport properties.

Phonons in bulk Si have a mean free path on the order of ~100 nm while the electron's mean free path is much shorter, on the order of ~ 1 nm[5]. Defects can be introduced so that phonons are scattered preferentially over electrons and holes. This could be accomplished by periodically removing parts of the Si material, so that there exists a Si/air interface. These defects are essentially an array of holes etched into the Si that have a diameter of > 10 nm and periodicity > 10 nm (Figure 5-4). The size of the holes will scatter phonons but should not significantly interfere with the electronic

transport. This will allow a reduced thermal conductivity from bulk Si while maintaining bulk electrical conductivity.

The structure in Figure 5-4 is constructed using SNAP and reactive ion etching (RIE)[6]. First, the SNAP process is performed on a degenerately doped SOI substrate with a top Si layer thickness ~20 nm. Next, another round of SNAP is performed orthogonally on top o the first set of SNAP nanowires. The nanowires in both SNAP rounds are made from Pt. Next, the RIE is used to etch into the top Si layer through the SNAP on SNAP pattern using a plasma of $CF_4$. Finally, the structure is placed in aqua regia to remove the Pt. The resulting structure is an array of holes etched into a thin Si film (Figure 5-4). The holes have a diameter and periodicity determined by the SNAP superlattice. The structure in Figure 5-4 will soon be electrically tested.
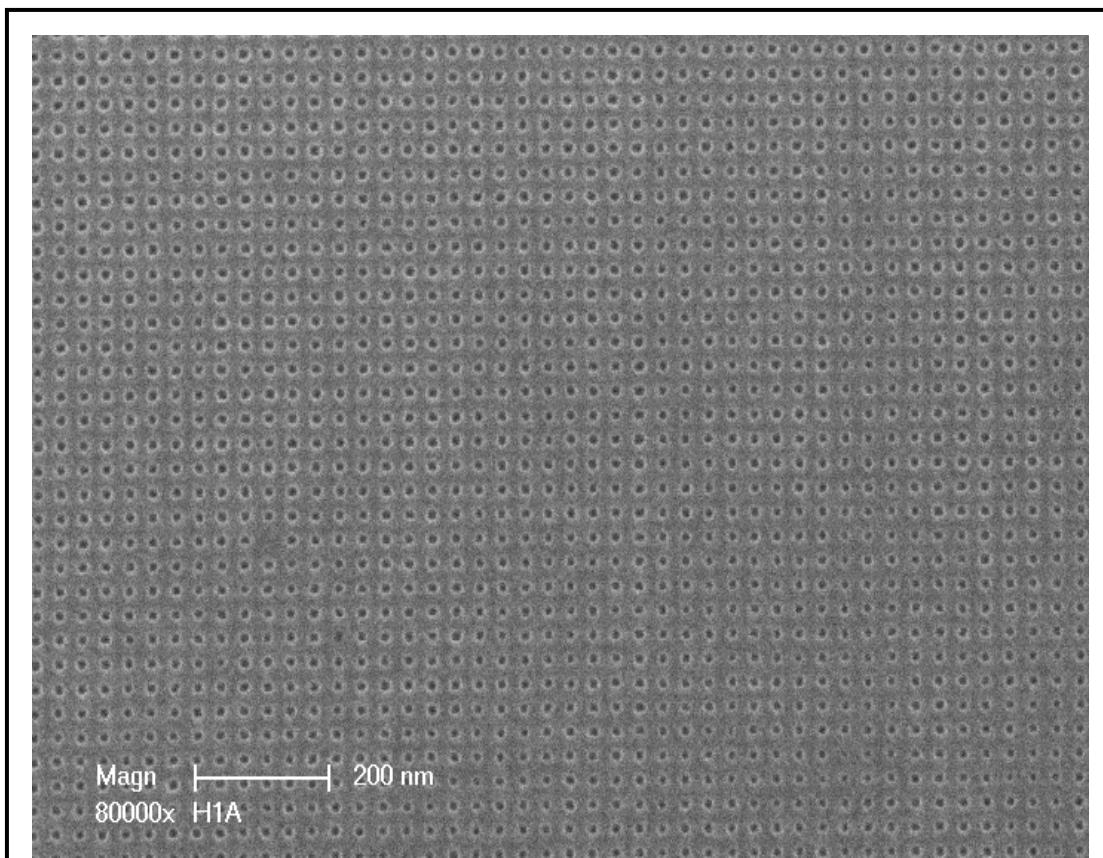


**Figure 5-4. A 20 nm thick degenerately doped Si film with 20 nm diameter holes etched into its surface.** This structure may provide the ability to scatter phonons while keeping the electrical transport properties bulk-like. Improvements in the figure of merit, **ZT**, should be possible.

## 5.3    References

1.  Humphrey, T.E. & Linke, H. Reversible thermoelectric nanomaterials. *Physical Review Letters* **94**, 096601 (2005).
2.  Kittel, C. & Kroemer, H., *Thermal Physics* (W.H. Freeman, New York, NY 1980).
3.  Humphrey, T.E. & H. Linke, Reversible Quantum Brownian Heat Engines for Electrons. *Physical Review Letters* **89**, 116801 (2002).
4.  Rowe, D.M. Ed., *CRC Handbook of Thermoelectrics* (CRC Press, Boca Raton, Fl 1995).
5.  Majumdar, A. Enhanced: Thermoelectricity in semiconductor nanostructures. *Science* **303**, 777-778 (2004).
6.  Wang, D., Bunimovich Y., Boukai, A. & Heath, J.R.  Two-dimensional, single crystal nanowire circuits.  *Small* **3**, 2043-2047 (2007).